

THIS WEEK

EDITORIALS

CLIMATE Global emissions deal depends on reliable data **p.288**

WORLD VIEW The big bad wolf – hunting politics in Sweden **p.289**

CAMOUFLAGE Coral-reef fish eat to conceal themselves using the chemicals **p.290**



Keep asking the question

Scientists must push to preserve a small part of a large US survey that provides essential information on the ever-changing scientific workforce.

Some 3.5 million households in the United States receive a survey each year with six dozen or so questions from the US Census Bureau. Among enquiries about occupation, income, household plumbing, commute times, ethnicity and more is 'Person Question 12', which asks university graduates what subject they studied. This census question, along with six others, may be dropped from future surveys as part of a push to streamline federal data collection. That would be a big mistake.

Data from question 12 are used by several studies of higher education, to assess, for example, how degree subjects correlate with unemployment and earnings. If the question is dropped, that information will be lost — or produced only at greater cost. *Nature's* readers can help to make sure that does not happen.

The question features on the American Community Survey, an ongoing mandatory survey launched in 2005 to provide timelier data than the more-intense decadal countrywide census. Faced with criticism from some legislators that the annual survey is a public imposition, officials reviewed all its questions to see how much time they required, how difficult or sensitive respondents found them, and how federal agencies used the data.

Asking about degrees posed a minimal burden on respondents, the review concluded. But the question was also deemed to be one of a few not required by statute or by regulatory agencies. (Alongside, for example, a question that asks whether US citizens have a medical facility on their property.) So, it faces the chop.

Why should it be kept? Difficult times for scientists make such data more important than ever. On 10 December, the National Academies released a long-awaited report on the postdoctoral experience. It decried the increasing fraction of PhD-holders taking these positions by default, and on academia's still-increasing treatment of postdocs as cheap labour rather than as trainees. Two weeks ago, this journal described two reports on the plight of postdocs and freshly minted science PhD graduates in the United States and the United Kingdom (see *Nature* 516, 7–8; 2014). Both reached similar conclusions: although an academic career is still presented as the default path, only a tiny minority (perhaps less than 5%) of new science PhDs will go on to permanent academic research positions. These reports stressed the need for more data to keep track of scientific (and non-scientific) careers.

The information that is available on the US situation is most conveniently presented in regular reports produced by the National Science Foundation (NSF). These include the biennial Science and Engineering Indicators and statistics about the participation and attainment of women, under-represented minorities, immigrants and disabled people. The reports reveal trends and disparities, such as the continuing dearth of women in computational science. They also aid international studies of the scientific workforce.

Congressional mandates demand that the NSF produces such reports. Before question 12 was introduced by the census bureau, the

agency had to carry out its own survey to acquire the information that the responses provide. It cost US\$17 million in 2003. Today, that effort would be even more costly and less effective. The workforce is increasingly mobile. People in scientific careers shift jobs so frequently that workforce scholars now refer to career pathways instead of pipelines. Building a sampling pool from the decennial census data would miss the hordes of people moving in and out of relevant fields as well as in and out of the country.

Scientists across the world are starting to realize the power and value of increasing efforts to study and foster its workforce. University offices and funding agencies are discussing how best to track the career paths of their graduate students and postdocs. The aim is to identify, promote and even create viable career paths outside the conventional system. To do so, these institutions need benchmarks — benchmarks that question 12 enables. Scientists and engineers are a rare population in statistical terms, and that means that less-intense population surveys are not big enough to get appropriate samples.

On a survey estimated to take 40 minutes to complete, question 12 requires only nine seconds. There is little to gain in its elimination and much to lose. Scientists and their allies should not only argue to retain the question, but also that the census bureau should recognize it as legally required in light of the NSF's mandates. The call for public comments on its removal ends on 30 December (see go.nature.com/ceqkkl). A robust response could encourage the administration to keep it. ■

"On a survey estimated to take 40 minutes to complete, question 12 requires only nine seconds."

Spin cycle

Pressures in all stages of the news-making process can lead to hype in science reporting.

It has become popular for people to receive, on landmark birthdays, a copy of a daily newspaper from the day of their birth. Someone born today, should they receive such a present in the future, may well wonder what on Earth they have in their hands.

The death of the printed daily paper has been much discussed. But the life of the printed daily paper is a curious thing, too: an entire existence predicated on the lie that the world has changed so much since the previous day that readers must pay for an instant briefing that they can hold in their hands. The same applies the following day, the day after that and so on.

The Internet has changed all of that, sometimes for the better and sometimes not. Yet one cultural legacy of the print-news world still rules: competition. Print readers were the ultimate consumers. Newspapers would compete for their patronage, and to make that happen, newspaper editors would make reporters compete for available space. Reporters would compete with rivals for stories. And anyone with a good story to tell had to compete with a thousand other people to get through to the reporter. The entire news-publishing business was an ever-decreasing circle, with someone on each step in the chain desperate to give the people on the next step exactly what they wanted.

What they all wanted, of course, was a good story — or more accurately, a better story than the other source, reporter, editor or newspaper was offering. Hence, routine speeches by politicians are often described as the most important of their careers, football matches with little at stake are ‘must-win’ and house prices are perpetually poised between collapse and meteoric rise. Good stories, naturally, are open to a little exaggeration; and a little more at the next step and so on. Newsroom culture demands that the most common phrase exchanged is not “Is this true?” but “Can we say this?”

Here comes the science bit. The reason that any of this matters to *Nature* is that science stories in the news, or more precisely, health and medical-science stories, are known to influence the behaviour of the people who read them. Together with the collective responsibility that many scientists feel for the way that research is communicated in the media (a responsibility that, say, estate agents seem to lack), this makes media coverage of research an important and much-scrutinized topic.

A study that has been heavily discussed over the past week or so focuses on the bottom step in the news chain described above: the information that universities give to reporters about published research (P. Sumner *et al.* *Br. Med. J.* **349**, g7015; 2014). The details appear on page 291 of this issue, but can be summarized as follows:

exaggeration in press reports of published medical-research papers is also present in press releases sent out by universities to promote those papers.

To conflate, briefly, correlation and causation (which the study counts as exaggeration), it seems that blame for media hype of medical research can be placed as firmly at the door of university press offices as on the headline-hungry keyboards of journalists.

“There is a demand for straight, less-conventional ‘news’ about science.”

Some journalists have nobly resisted the temptation to pass the blame in this way, and insisted that their profession must do more to check the claims made by others before handing them on. Others have called for stricter controls on what universities say, and for scientists who have their work promoted to be held accountable. These are all sensible ideas, and *Nature* fully supports the idea that researchers should work closely with those who write and circulate press releases on their behalf.

Exaggeration will persist in the news cycle only if it benefits all those involved — from the scientists who can count press coverage as ‘impact’ to the reporters who bag another high-profile byline and the approving comments of their bosses.

But will it persist? Coming back to the description of newsroom culture, “Can we say this?” is itself giving way to “What else can we say?” as elastic electronic boundaries of news websites replace physical page budgets. The rise (and mass readership) of specialist blogs shows that there is a demand for straight, less-conventional ‘news’ about science. The implicit benefit of exaggeration — to help stories to squeeze through the next stage in the news process — is weakening.

The study suggests as much — there was no link between the amount of exaggeration in a press release and the media coverage that it received. The truth, in other words, does not have to hurt. ■

Honest brokers

Climate negotiations in Lima stumbled on transparency, but there is time to adjust.

The main task for negotiators at the United Nations climate talks in Lima last week was simple: lay out the rules for the emissions pledges that countries will submit over the next six months. Countries had already agreed to put forth plans, each according to its own needs, capabilities and circumstances, and were riding a small wave of optimism after the surprise announcement in the lead-up to the talks that China and the United States had agreed to cut their emissions. The question was how to register and interpret these commitments going into the headline summit in Paris next year.

It is hard to overstate the simplicity of this task, especially relative to the magnitude of the challenge at hand. And yet negotiators went into double overtime fighting old fights, and walked away with something that bears a clear resemblance to nothing.

Negotiators had various options on the table, ranging from a generic registry of commitments to a formal review process in which countries would be expected to provide the relevant data and then defend the adequacy of their pledges. But after days of bickering about what should be required of whom — led by China, which opposed the reviews — they wound up with a text that requires little of anybody.

The final system must allow everybody to evaluate all national commitments and track their progress over time. A treaty that formalizes such an approach would give all countries confidence that their investments are not in vain.

Sure, nations are beginning to take action, but it is the cumulative

carbon emissions that matter. The end goal is a world with essentially zero emissions. That is not possible unless all countries play ball. We are in the middle of a trust-building exercise, and the first step is transparency.

One sticking point is that national commitments can (and will) be assessed in various ways. Wealthy countries will measure actual reductions in emissions; rapidly developing countries might opt for reductions from forecast growth. But commitments can also be assessed in terms of cost, either absolute or relative to economic activity, and even on technical capacity for the poorest nations. Both carbon emissions and investments can be assessed relative to population and per-capita income to get at the question of equity, which is at the heart of most disputes in the climate negotiations. All of these measures are legitimate, and academics are already busy with such analyses. But they all depend on one thing: information, which is what was dropped from the Lima agreement.

Some countries are likely to provide the relevant evidence to bolster their cases, but this process must be streamlined and must be required of every country. Governments, scientists and environmentalists will fill in any gaps as best they can over the coming year, but the challenge will only grow. Next year’s pledges will probably fall well short of what is needed to prevent the worst impacts of global warming, so commitments will need to be reviewed and updated regularly. Once governments can demonstrate progress, the plan is for them to initiate a virtuous cycle in which better policies and cheaper technologies help to push emissions ever lower.

This will only work, however, if governments can be held accountable and independent analysis can identify which policies are working — and which are not. And to do that, the world will need solid data and robust assessments. Simple or not, the treaty to be signed in Paris should recognize as much. ■

➤ NATURE.COM
To comment online,
click on Editorials at:
go.nature.com/xhunqv



Challenge the abuse of science in setting policy

The misuse of wolf research by Swedish politicians should be a warning to all biodiversity scientists, says Guillaume Chapron.

As the northern winter takes hold, Sweden is preparing for its largest ever wolf hunt. The country has been trying to hunt significant numbers of the animals for years — in the face of a European law that lists them as a strictly protected species — and it looks as though it will now succeed.

The situation is particularly alarming for me because the government has incorrectly used my academic research to make its case that the wolf population has recovered.

Political tensions over recovering populations of large carnivores are common in Europe. But the wolf issue in Sweden is unique because scientific knowledge and how it is interpreted have become central to justifying hunting. The conservative Swedish government has been playing with scientific findings for political reasons. It has claimed that its decisions are supported by the research it asked me to produce — but they are not — and it has cherry-picked others' findings. The situation is at odds with the popular view of the supposed respect that Nordic countries have for evidence-based environmental sustainability.

There are about 400 wolves in central Sweden and the population is heavily inbred: all the wolves are descended from a handful of animals that have arrived from Finland since the 1980s. The European Habitats Directive, which protects the wolf (*Canis lupus*), does allow for limited culling to prevent serious damage to livestock. But there is consistent political pressure to reduce wolf numbers further. For example, hunters complain that every year the wolves kill a few hunting dogs, which run free as their owners target moose.

Beginning in 2010, the Swedish government claimed that annual wolf hunts, which aimed to slash numbers to 210 animals, would persuade hunters to support plans to import unrelated wolves from Finland or Russia and make the population more genetically diverse. However, although the hunt went ahead, disease fears scuppered translocation of the foreign animals.

In following years, the annual hunts faced various legal challenges, and by 2013, the government had a new scientific justification. It said that hunting was the single most effective way to immediately solve the wolf population's genetic problems. Shooting the most inbred wolves, the government pointed out, would at a stroke decrease the inbreeding coefficient of the population.

I told the Swedish authorities that this was a deliberately short-sighted idea because the only way to decrease inbreeding in the long run is to bring in new genes. A complementary and risky proposal to translocate captive-bred pups into wild litters failed too. That year's hunt began anyway, but was halted by the Swedish courts.

Despite vitriolic letters from the European Commission calling on Sweden to make sure that the wolf population reaches Favourable Conservation Status (FCS) — a mandatory benchmark of a recovered and thriving population — the Swedish government did not give up. Late last year, it ditched its genetic concerns — the only reason it had given to support the hunt just twelve months before — and simply declared that the wolves had reached FCS.

This is where my research was misused. In 2012, the Swedish government gave me 30 days to prepare a population viability analysis of the wolves. This is a demographic measure of how close the population is to extinction, and crucially, is a separate measure from FCS, which relates to recovery. To avoid misinterpretation of my work, which excluded genetic aspects, I made sure to write on multiple occasions in the report

that it could not be used to estimate FCS. Several reviewers of the report also stressed this point.

Yet the government still misused my report to claim that the wolf population in Sweden had reached FCS, as a cover to permit further hunting.

As preparations for this year's hunt continue, legal protection for the wolves is harder to find. Realizing that scientific evidence will be a constant obstacle, the government has changed the law to effectively make large-carnivore-hunting decisions exempt from legal challenge. Furthermore, it opposed mention of the need for research on FCS in a forthcoming European action plan for large carnivores, arguing that the Swedish parliament had voted on FCS so there was no need for such research. When my project 'Claws & Laws', which is aimed at exploring FCS, was funded by the independent Swedish Environmental Protection

Agency, some politicians made known their uneasiness with the work.

I am concerned that Sweden's misuse of my research and its flouting of European regulations will set a dangerous precedent in biodiversity conservation. The distortion of science has been very subtle and technical in this case, and the wolves will not be eradicated, but it is important to highlight because it may be the first of many examples. Preserving biodiversity can generate conflict because it places limits on development, traditions and other human activities. Ecological science will probably have a more important role in these disputes in the future.

With increasing calls to make policy science-based, political abuse is likely to become more common. Even if it damages their careers, and makes their names toxic, academics must be prepared to identify the unethical use of scientific knowledge and expose such abuse by politicians. ■

Guillaume Chapron is associate professor at the Swedish University of Agricultural Sciences in Riddarhyttan, Sweden.
e-mail: guillaume.chapron@slu.se Twitter: @CarnivoreSci

**PRESERVING
BIODIVERSITY
CAN GENERATE
CONFLICT
BECAUSE IT PLACES
LIMITS ON
HUMAN
ACTIVITIES.**

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/sxy22u

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

PALAEOCLIMATE

Lake cores support legend of typhoons

Geoscientists have found possible evidence of two typhoons that, according to Japanese legend, wiped out invading Mongol fleets in the years 1274 and 1281.

Jon Woodruff of the University of Massachusetts, Amherst, and his colleagues collected a 2,000-year-old sediment record from a coastal lake on Japan's Kyushu island, where the Mongol attack was aimed. The cores contain flood deposits that show two instances of flooding in the late thirteenth century, which may have come from the pair of 'Kamikaze' typhoons.

Such storms could have been more common at the time, thanks to the presence of an El Niño, which causes changes in temperature and precipitation worldwide.

Geology <http://doi.org/xqp> (2014)

METEOROLOGY

Lopsided hail hits harder

Hail storms can cause billions of dollars' worth of damage, but until now scientists have known little about the precise mass and shape of hail. A study has found that hailstones that are not perfectly spherical can sometimes travel faster and hit objects with greater force than



spherical hailstones, potentially causing severe damage to homes and cars (pictured).

Andrew Heymsfield of the National Center for Atmospheric Research in Boulder, Colorado, and his colleagues measured nearly 2,300 hailstones that fell across the US Great Plains between 2012 and 2014. Most hailstones were smaller than 3 centimetres in diameter, but those that were bigger tended to be more lopsided than the smaller ones. Calculations suggest that the non-spherical hailstones occasionally hit objects with greater force than would be expected if they were round.

The findings could help

to improve weather and hail-damage predictions, the authors say.

Geophys. Res. Lett. <http://doi.org/xqq> (2014)

MOLECULAR EVOLUTION

How bacteria and host fight for iron

A study of primate and bacterial proteins involved in capturing iron from the blood has revealed an evolutionary arms race in the battle over this important nutrient.

Matthew Barber and Nels Elde at the University of Utah in Salt Lake City focused on transferrin, a protein that

they exposed species of coral-inhabiting crabs to the odour of fish that ate that same coral species, the crabs preferred those fish to animals that ate another coral. This suggests that the filefish's diet influences its scent.

Moreover, a filefish predator, cod, had difficulty detecting the fish when they were near the coral on which they were fed. The study provides the first evidence for chemical-based camouflage in a vertebrate.

Proc. R. Soc. B <http://dx.doi.org/10.1098/rspb.2014.1887> (2015)



ANIMAL BEHAVIOUR

Fish adopt chemical camouflage

A coral-reef fish can match its scent to the odour of the surrounding reef, masking itself from predators.

Harlequin filefish (*Oxymonacanthus longirostris*; pictured) live around reefs in the Pacific and Indian oceans and feed on particular species of coral. A team led by Rohan Brooker at James Cook University in Queensland, Australia, tested whether aquarium-dwelling fish conceal themselves by emitting a scent that is similar to those generated by the corals that they consume. The authors found that when

transports iron from the blood into cells. Pathogenic bacteria compete for this mineral by using their own protein, called TbpA, to bind transferrin. The researchers sequenced transferrin from 21 primate species to trace its 40-million-year evolutionary history, and tested the molecules' interactions with TbpA from two common human pathogens. They found specific amino-acid changes in a rapidly evolving region of transferrin that prevent TbpA from binding to it.

They also pinpointed transferrin-binding sites in TbpA that are genetically diversifying under selection,

DAVID FLEETHAM/OCEANWIDEIMAGES.COM

MIKE DREW/QMI/CORBIS

showing how competition for a nutrient can drive primate and pathogen evolution. *Science* 346, 1362–1366 (2014)

ANTHROPOLOGY

How a Maya city rose and fell

One of the major Maya cities thrived in a tropical forest by using sophisticated agricultural, forestry and water-management techniques.

David Lentz at the University of Cincinnati in Ohio and his colleagues surveyed modern forests at the site of Tikal in Guatemala, which was a bustling city roughly 1,400 years ago. By analysing archaeological plant and soil specimens, the authors concluded that the people of Tikal intensively farmed the land, using irrigation and terraces, for example. They also developed a complex system for collecting and distributing rainwater.

However, by around AD 850, as drought set in, the Tikal systems could not keep up with the growing population, probably leading to the demise of the great city.

Proc. Natl Acad. Sci. USA
<http://dx.doi.org/10.1073/pnas.1408631111> (2014)

PHYSICS

Record-breaking electron boost

Physicists have used lasers to increase the amount of energy that electrons gain per metre by more than two orders of magnitude compared with traditional accelerators.

Conventional colliders can accelerate particles to much greater energies, but over many kilometres. Wim Leemans at Lawrence Berkeley National Laboratory in California and his colleagues used extremely intense laser pulses and an ionized gas to boost electrons over much smaller distances.

By guiding the pulses through channels in the plasma, the researchers

generated strong electric fields that increased injected electrons to 4.2 gigaelectronvolts — the highest energy ever achieved in a laser-based system — over just 9 centimetres.

The authors say that the technique could be used to make smaller high-energy linear accelerators, and to create table-top systems that use X-rays emitted by electrons to probe materials. *Phys. Rev. Lett.* 113, 245002 (2014)

CHEMISTRY

Painkillers made in minutes

Ibuprofen can be produced in minutes by mixing reagents as they flow through a series of connecting tubes.

Synthesizing a substance in a continuous-flow process offers more control over reactions and allows less solvent to be used than batch production in flasks. But solid by-products can accumulate, blocking the flow. By varying the width of connecting tubes and using specially designed pumps, David Snead and Timothy Jamison at the Massachusetts Institute of Technology in Cambridge built an apparatus that produces ibuprofen in three minutes with minimal fouling. The five-stage process has three reactions — one of the most complex applications of flow chemistry yet.

A wide variety of other drugs could be synthesized in this way, the authors say.

Angew. Chem. Int. Edn <http://doi.org/10.1002/anie.201402444> (2014)

AGRICULTURAL ECOLOGY

Pesticide moves up food chain

An insecticide banned in some areas for its effect on bees not only fails to kill certain pests, but also harms the predators that feed on them.

Neonicotinoid insecticides are used on many crops, including soya-bean plants, on which pest slugs (*Derocera reticulatum*) feed. Margaret

SOCIAL SELECTION

Popular articles
on social media

When press releases go bad

Researchers love to blame the news media when reports about science are misleading or even wrong. But a study making the rounds online suggests that much of the hype and misinformation about health-related research in the news has its roots in university press releases — which are often approved in advance by the researchers themselves. “Academics should be accountable for the wild exaggerations in press releases of their studies,” tweeted Catherine Collins, a dietitian who works for the National Health Service in London. But some say that others are to blame. “Exaggerated academic hype leads to bad news stories. Why don’t reporters do their jobs?” tweeted Steve Usdin, editor and co-host of *BioCentury This Week*, a US public-affairs show covering the biopharma industry. *Br. Med. J.* 349, g7015 (2014)



Based on data from altmetric.com. Altmetric is supported by Macmillan Science and Education, which owns Nature Publishing Group.

➔ **NATURE.COM**
For more on popular papers:
go.nature.com/czhask

ASTRONOMY

Exoplanet seen from Earth

Using a modest-sized ground-based telescope, astronomers have spotted a planet twice the size of Earth passing in front of its host star.

Researchers typically study planets outside the Solar System using space telescopes or much larger telescopes on Earth, but studies with space telescopes are expensive and access to large facilities on the ground is limited. A team led by Ernst de Mooij, now at Queen's University Belfast, UK, used a smaller telescope in La Palma, Spain, to investigate the Sun-like star 55 Cancri. The scientists' size measurements of one of the star's known planets were similar to those obtained using orbiting telescopes.

Such ground-based studies can complement those using space telescopes, the authors say.

Astrophys. J. Lett. 797, L21 (2014)

➔ **NATURE.COM**
For the latest research published by Nature visit:
www.nature.com/latestresearch



Douglas at the Pennsylvania State University in University Park and her colleagues exposed the slugs in the lab to soya-bean plants grown from seeds coated with the neonicotinoid thiamethoxam. They found that the slugs were unaffected, but that more than 60% of ground beetles (*Chlaenius tricolor*, pictured), which feed on the slugs, died or suffered impairments such as paralysis. In field studies, thiamethoxam also lowered the number of predators on slugs, and reduced soya-bean yields by 5%.

The results indicate unintended indirect effects of neonicotinoids on non-target species in addition to known direct effects, the authors say. *J. Appl. Ecol.* <http://doi.org/10.1111/1365-2656.12444> (2014)

SEVEN DAYS

The news in brief

RESEARCH

Study stopped

The US National Institutes of Health (NIH) has cancelled plans for a multi-decade study of children's health, agency director Francis Collins announced on 12 December. Commissioned by the US Congress in 2000, the National Children's Study was to assess how physical, chemical, biological and psychosocial factors affected 100,000 children from birth to the age of 21. The NIH has spent US\$1.2 billion on the effort and enrolled roughly 5,700 children in a pilot study at 40 centres. But the project has been delayed by scientific disagreements and management problems. See go.nature.com/i8xwyy for more.

Collider comeback

CERN, Europe's particle-physics laboratory near Geneva, Switzerland, confirmed on 12 December that the Large Hadron Collider is on track to restart in March 2015. The planned reboot follows a two-year shutdown, during which the accelerator and detectors have been upgraded to work at a record collision energy of 13 trillion electronvolts. The machine is now close to being cooled to its operating temperature of 1.9 kelvin, and on 9 December, the magnets of one sector were successfully powered to operating levels.

Microbial menace

Left unchecked, antimicrobial resistance could cost the world up to US\$100 trillion by 2050 and cause 10 million deaths per year, according to a panel commissioned by the UK government. The panel, chaired by economist Jim O'Neill, released its first report on 11 December. The projections show how predicted rises in resistance are likely to affect health, the



RODRIGO ABD/AP

Greenpeace harms archaeological relic

Peruvian government officials said on 9 December that they will pursue legal action against Greenpeace activists who damaged the site of the country's famous Nazca lines by installing a campaign message next to the ancient etched figure of a hummingbird (pictured). Members of the environmental group had sought to promote renewable energy with large cloth letters visible from the air during the latest round of United

Nations climate negotiations in the capital, Lima. But deputy culture minister Luis Jaime Castillo said that the activists had entered a strictly prohibited area of the UNESCO World Heritage Site, and had irreparably disturbed patterns in the dirt. The government is seeking to detain the activists in Peru, said Castillo, and charge them with attacking archaeological monuments — punishable by up to six years in prison.

labour force and economic production. They probably underestimate the threat, the authors say, because the study examined only drug-resistant bacteria and public-health issues for which data were readily available.

PubPeer fights back

The team behind PubPeer, a website for discussing scientific articles, filed a legal motion on 10 December to quash a subpoena by cancer researcher Fazlul Sarkar at Wayne State University in Detroit, Michigan. Sarkar says that anonymous comments about his work on PubPeer are defamatory; the University of Mississippi in Oxford withdrew a job offer to him after seeing

the comments. Sarkar has subpoenaed PubPeer to reveal identifying information about the commenters (see *Nature* <http://doi.org/w68>; 2014). PubPeer's motion argues that the comments are not defamatory, and that the subpoena jeopardizes the free speech needed for scientific progress.

POLICY

US budget

The US Senate passed a US\$1.1-trillion spending bill on 13 December, which would boost funding for NASA and the National Science Foundation in fiscal year 2015. It would also give \$5.4 billion in aid and research

funds for the Ebola epidemic in West Africa, but would raise overall funding for the National Institutes of Health by only about 0.5%. President Barack Obama is expected to sign the bill into law, finalizing the budget for US agencies until 30 September 2015. See go.nature.com/tkm71a for more.

Nuclear power

Russian President Vladimir Putin and Indian Prime Minister Narendra Modi announced on 11 December a raft of oil, defence and nuclear agreements, including a plan for Russia's state-owned nuclear energy corporation to supply at least 12 new nuclear reactors to India over the next

ADEEL HALIM/BLOOMBERG/GETTY



20 years. India's government has pushed to expand its nuclear-power capacity, despite deep public opposition (see *Nature* <http://doi.org/ckcr86>; 2011). Six of the new reactors will be at the Kudankulam power plant (pictured) near India's southern tip, which already hosts two Russian-built reactors (see *Nature* 499, 258–259; 2013).

Climate deal

Two weeks of climate talks in Lima have produced a road map for an international climate treaty to be negotiated in Paris next year. The deal, announced on 14 December, lays out basic rules for how countries should formulate and submit pledges to reduce greenhouse-gas emissions. Those pledges are expected in the first half of next year. Faced with opposition from nations including China, negotiators abandoned language that

would have established formal reviews for climate pledges and would have required countries to submit technical data to help to evaluate those pledges. See go.nature.com/bcxgna for more.

EU budget

On 8 December, European Union (EU) governments reached a last-minute provisional deal with members of the European Parliament on a €141.2-billion (US\$176-billion) budget for 2015. Parliamentarians secured an extra €430 million on top of a budget proposal that lawmakers had rejected last month (see *Nature* <http://doi.org/xqf>; 2014), including €45 million more next year for the Horizon 2020 research programme. Governments have committed to provide an additional €4.8 billion to reduce the €23.4-billion backlog of unpaid bills. The

deal is expected to be signed by the 28 member governments and to pass a full Parliament vote next week.

Research fund

Australia's government has abandoned plans to finance a Aus\$20-billion (US\$16.5-million) medical-research fund by charging people to visit their family doctors. After opposition from the public and medical professionals, Prime Minister Tony Abbott said on 9 December that the Medical Research Future Fund would go ahead, but the compulsory Aus\$7 charge would not. Instead, the research funding will come from savings in the health-care budget, such as a reduction in payments to doctors for patient visits — a cost that doctors could choose to impose on patients, although pensioners, children and some others are exempt from the charge.

FUNDING

Faster funding

The California Institute for Regenerative Medicine (CIRM) in San Francisco on 11 December approved a US\$50-million plan to overhaul its research funding mechanisms. Starting on 1 January 2015, 'CIRM 2.0' will aim to fund successful applications within four

months of submission — a process that in the past could take up to two years. The plan, which would also give researchers more chances each year to apply for funds, is designed to attract clinical-stage research that is ready to start within 45 days of approval.

Ebola vaccines

Gavi, the vaccine alliance based in Geneva, Switzerland, announced on 11 December that it will pledge up to US\$300 million to buy up to 12 million courses of Ebola vaccines to immunize at-risk populations. Gavi is awaiting recommendations on a safe and effective vaccine from the World Health Organization. Clinical trials are currently under way, including one of a vaccine developed by Merck and NewLink that researchers announced they had suspended on 11 December after four patients complained of joint pains. In addition to funds for vaccine procurement, Gavi committed up to an additional \$90 million to help to introduce vaccines and to rebuild health systems in countries affected by Ebola.

BUSINESS

Pharma leader

Beginning in early 2015, geneticist David Altshuler will join Vertex Pharmaceuticals in Boston, Massachusetts, as chief scientific officer and executive vice-president for global research, the company announced on 15 December. Altshuler, who was a founding member of the Broad Institute in Cambridge, Massachusetts, will lead the company's drug-discovery efforts and oversee research at five sites in the United States, Canada and Europe. He currently holds faculty positions at Harvard Medical School and the Massachusetts Institute of Technology, and practises medicine at Massachusetts General Hospital in Boston.

➔ **NATURE.COM**

For daily news updates see:
www.nature.com/news

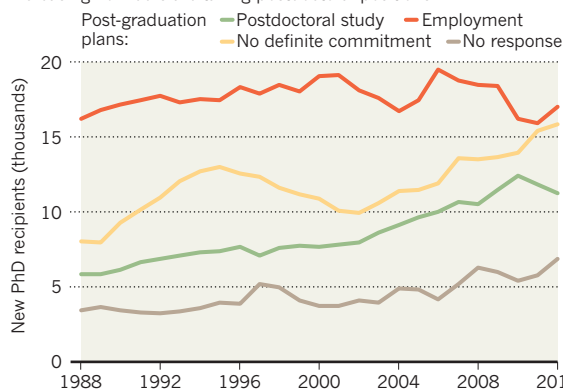
SOURCE: US NATL ACADEMIES

TREND WATCH

The number of people receiving PhDs in the United States, especially those in engineering and the biomedical sciences, has outpaced job opportunities outside academia in recent years, according to a report by the US National Academies in Washington DC. Fewer graduates have job commitments than in the past, and more are taking postdoc positions (see chart). The report recommends the creation of fixed-term postdoc positions to help prevent graduates from working for long periods on low salaries.

POSTDOC PROLIFERATION

Since 1988, the number of new PhD recipients in the United States reporting post-graduation employment has stagnated, while increasing numbers are taking postdoctoral positions.



NEWS IN FOCUS

EVOLUTION Flock of geneticists redraws bird family tree **p.297**

GEOPOLITICS Tempers explode at Russian science meeting **p.298**

2014 IN REVIEW Comets brought triumph but stem cells brought tragedy **p.300**

NATURE'S 10 Ten people who made a difference to science this year **p.311**



TOMMY TRENCHARD



A burial team inter a person who had died in an Ebola holding centre in Makeni, Sierra Leone.

DISPATCHES

Ebola threatens a way of life

A report from Sierra Leone examines the cultural struggle to eradicate the virus.

**BY ERIKA CHECK HAYDEN,
BOMBALI DISTRICT, SIERRA LEONE**

Since September, the Ebola virus has stalked the villages and towns along the Kamakwie–Makeni Road, a rutted, red-dirt track that serves as the main artery for a string of villages in the western part of Sierra Leone's Bombali District.

Yeli Sanda, a village just a few kilometres outside the district's capital city of Makeni, was the first place to be hit. Over the following months, more than 40 people in the settlement of about 700 became infected; 22 died. In November, the virus infected a woman in Tambiama, about 11 km up the road. A friend who visited her acquired the virus and carried

it another 1.5 km to the village of Mayata. She and at least five others there have died.

But just a few hundred metres from Yeli Sanda, the village of Yoni has not seen a single case of Ebola. As soon as the village chief learned that Ebola had struck, he forbade his citizens from visiting Yeli Sanda or attending burials of its residents. His swift action has kept the 100 people of Yoni healthy while other communities have been devastated.

Public-health officials and local leaders who have volunteered in the Ebola fight say that Yoni's experience is instructive: to banish the virus from the Kamakwie–Makeni Road, which runs for 150 km

from Sierra Leone's fourth-largest city to the Guinean border (see 'On the road'), they will have to convince people to abandon some long-held beliefs and customs.

"If we don't get cooperation from traditional institutions, we could spend a long time chasing this, village to village, all the way to the border," says Adam Goguen, the registrar at the University of Makeni and a resident of Yoni.

His village looks like any other in the district: a clutch of houses fronting dusty yards, backed by small farming plots cleared out of the scrubby forest. But the exceptional behaviour of its residents — their willingness to cut social ties and abandon cherished traditions — has so far kept it safe. As difficult as it is, public-health officials say, changing behaviour is the key to ►

► **NATURE.COM**
Video diary from
Sierra Leone:
go.nature.com/phwqhd



ON THE ROAD

Since September, local officials have fought Ebola in villages along the Kamakwie-Makeni Road in Sierra Leone.

▶ stopping the Ebola outbreak that has ravaged West Africa for a year. “No more social life. No more business. No more travelling,” says Sorie Bundu Conteh, a disease-surveillance officer. “We need people not to see us as a threat,” he adds. But that is difficult when officials are asking people to restrict their lives so drastically for reasons that can be hard to understand.

In Yeli Sanda, communication problems began with the very words that local officials first used to talk about Ebola: there is no word for ‘virus’ in the tongues spoken in the villages along the Kamakwie-Makeni Road. Before the outbreak reached the area, Ebola educators there described the pathogen as a kind of tumbu, or maggot. When Ebola came to Yeli Sanda, a man searched through the blood of someone who had died from it, looking for the maggots. In doing so, he potentially exposed more people to the virus.

Yoni chief Pa Alpha Tarawalie understood the situation better than most. He says that he decided to cut his people off from Yeli Sanda on the basis of what he had heard on BBC radio broadcasts, and seen while delivering supplies to quarantined communities on the outskirts of Makeni. With accurate information and a witness’s appreciation of the disease’s devastating effects, he says, he was able to convince his people to sever ties with their closest neighbours.

The measure angered people in Yeli Sanda, who felt shunned, and even accused the Yoni

nurse of witchcraft when she correctly predicted that people who attended Ebola victims’ funerals would contract the disease themselves. But the rift between the two villages has started to heal now that Yeli Sanda’s outbreak has subsided. In November, Yeli Sanda’s chief was suspended for flouting disease-control rules.

SECRET BURIAL

One of the most effective ways to stop a disease such as Ebola is contact tracing — tracking down and quarantining everyone who may have had close contact with a person who shows symptoms. But when a prominent community member in Yeli Sanda became infected, traditions made contact tracing impossible. The man was a member of Gbangbani, one of several ‘secret’ societies in Sierra Leone that can have an important role in village life. The groups are far from secret in terms of membership — in many villages almost every adult belongs to one. But their proceedings and rituals, including their burial rites, are kept out of public view.

When the man died, his fellow members took his body to a remote village and kept to tradition by burying it under cover of darkness. Handling the body of someone who has died of Ebola puts

a person at great risk of contracting the disease — normally, anybody who did this would face weeks of quarantine. But nothing about the Gbangbani burial, from the participants to the location, was shared with public-health officials.

“We’ve spent years trying to change behaviours, for instance in the HIV epidemic,” says Peter Salama, global emergency coordinator for Ebola with the United Nations charity Unicef. In the case of Ebola, “we’ve got weeks or months”.

STAND-OFF IN MAYATA

Some customs, like the burial of the Gbangbani man, are specific to their cultures. Others, such as attending a neighbour’s funeral, are more universal. But they can be just as deadly. The woman who introduced Ebola to Mayata did nothing more unusual than visit an ill friend in a neighbouring town. With her and five other people in Mayata now dead, surviving relatives and neighbours are inclined to run. That has pulled them into conflicts with public-health workers attempting to enforce a quarantine.

In early December, Father Francis Sehdu Sesay, a dean of theology at the University of Makeni, found village elders in a stand-off with a Mayata woman who was the only survivor out of four adults in her household. She had fled for several days to an unknown location and had just returned to the village. But she refused to go into quarantine back at the house, where four orphaned children remained.

As Sesay and the elders pleaded with the woman, she stared straight ahead, her arms folded across her chest. She feared being sent to a distant district for treatment. Sesay explained that she would not be sent away — only watched for symptoms. And if she was diagnosed with Ebola, she could be treated at a local centre that had just opened.

The woman relented at last, and agreed to be quarantined in her home. Sesay climbed back into his vehicle and pulled out onto the Kamakwie-Makeni Road. He stared out of the window at the forest that stretches for kilometres in every direction, wondering where in the country’s vast interior the woman had fled to.

“What about inside, where vehicles and burial teams don’t go?” he asked. “People go to these villages, away from the road, and spread it there.” ■

The Pulitzer Center on Crisis Reporting provided support for this coverage.



**MORE
ONLINE**

TOP NEWS



Superconductors claimed to smash temperature record
go.nature.com/yabkhj

MORE NEWS

- Clinical trial for Ebola survivor serum kicks off go.nature.com/estpst
- Ocean floors full of plastic waste go.nature.com/hpdghr
- Cloud-seeding test over Wyoming’s mountains is inconclusive go.nature.com/dymife

NATURE PODCAST



‘Silly season’ special includes charades, a quiz and much more
nature.com/nature/podcast

PHIL DEGGER/ALAMY



Bird genomes are helping to reveal the relationships between species.

EVOLUTION

Bird family tree is in fine feather

Behind the most comprehensive tree of life lies a vast collaboration of like-minded researchers.

BY EWEN CALLAWAY

Evolutionary geneticist Tom Gilbert was sipping a coffee in Madrid five years ago when an idea hit him — literally. “A pigeon crapped on me,” he says, “and I thought to myself, ‘Huh, pigeons.’”

On 11 December, Gilbert, of the Natural History Museum of Denmark in Copenhagen, and dozens of his colleagues reported an evolutionary analysis of the genomes of 48 bird species (including pigeons), amounting to the most comprehensive genome study of any major branch of the tree of life¹. The results confirm a ‘big bang’ in bird diversity after dinosaurs went extinct, and settle long-standing questions on how different birds are related to each other.

On the same day, a consortium of researchers co-led by Gilbert published a further 18 bird-genome papers in various journals, on topics as diverse as the basis of birdsong, birds’ loss of teeth and the cold-weather adaptations of penguins (see avian.genomics.cn/en).

No one has ever used so much genome data from so many species to determine evolutionary relationships. Achieving this daunting task meant building a vast international collaboration that began, appropriately, with pigeons.

In 2010, Gilbert struck up a partnership with BGI, a sequencing powerhouse in Shenzhen, China, to map the first pigeon genome. The goals were to work out how different breeds relate to each other and the origins of their various traits. Gilbert met BGI genome scientist Guojie Zhang later that year and discovered that the BGI had sequenced several other bird genomes for a project led by neuroscientist Erich Jarvis of Duke University in Durham, North Carolina. The three researchers realized that, with a few more samples, they could get genomes from all branches of a group called Neoaves, which includes most modern birds except flightless species (such as ostriches and emus) and chickens, ducks and other fowl. “It struck me that there’s this pretty major unsolved question in avian evolutionary history, which is how do the different bird orders relate to each other?” Gilbert says.

No one had been able to determine which species split off first from the common ancestor of all Neoaves. Furthermore, study after study had thrown up different ways of mapping the evolutionary relationships between the subset of Neoaves

that exhibit vocal learning, a relatively rare trait that scientists see as analogous to human speech. Only entire genomes would reveal birds’ true evolutionary history, Gilbert and his colleagues surmised.

Gathering DNA samples was fairly straightforward. So, too, was sequencing the genomes, which BGI finished by summer 2011. But analysing the data and using them to build an evolutionary tree required another three years, new computational methods and 300 years of computing time. Hundreds of researchers asked if they could use the data, and the project swelled to 80 institutions in 20 countries; marathon Skype calls became a weekly fixture.

The results illuminate various aspects of bird biology, from neurophysiology to population genetics. In one effort, Jarvis and his co-workers discovered parallels between gene activity patterns in brain areas involved in birdsong and in human speech². Another effort dated the loss of teeth in birds to around 116 million years ago³. Yet another showed how inbreeding had shaped the genome of the crested ibis (*Nipponia nippon*) after a recovery programme brought its population up from seven individuals in the 1980s to hundreds now⁴.

The genomes also reveal the broad brushstrokes of the bird family tree. The results show that the first Neoaves species to peel off were ancestors of today’s doves, grebes and flamingos. The authors also conclude that vocal learning may have evolved independently in the ancestors of parrots, hummingbirds and songbirds, and that the ancestor of all land birds — which include eagles, woodpeckers, crows and parrots — was probably similar to a modern bird of prey or, as Gilbert puts it, “a mean-ass carnivore”.

The genomes also point to an explosion in diversity between 67 million and 50 million years ago, a period when non-bird dinosaurs are thought to have been wiped out by an asteroid impact. Mammals seem to have flourished then too, and both groups may have taken advantage of the niches that dinosaurs left behind.

Stephen Richards, a genomicist at Baylor College of Medicine in Houston, Texas, who is leading an effort to sequence 28 insect genomes, praises the team’s decision to systematically select bird species so that one from each taxonomic order was represented, rather than picking scientists’ favourite species. “It’s a foundational work for the next century of biological work into birds,” he says. “We need this revolution across all of biology.”

Gilbert, meanwhile, is a convert to super-sized projects that bring together multiple labs. No single group can do all the work to answer other questions he wants to tackle, such as the evolution of domestic crops. “I don’t spend all my time looking at hummingbirds,” he says. Or pigeons, for that matter. ■

NATURE.COM
See the latest bird
tree of life at:
go.nature.com/9num49

1. Jarvis, E. D. *et al. Science* **346**, 1320–1331 (2014).
2. Pfenning, A. R. *et al. Science* <http://doi.org/xqh> (2014).
3. Meredith, R. W. *et al. Science* <http://doi.org/xqn> (2014).
4. Li, S. *et al. Genome Biol.* **15**, 557 (2014).



Protesters proclaim the death of Russian science at a 2013 rally against education reforms in Moscow.

POLITICS

Putin's Russia divides scientists

Are geopolitical tensions destroying important links with the West, or can Russian research go it alone?

BY QUIRIN SCHIERMEIER, ST PETERSBURG

Even before the slide bearing a portrait of Joseph Stalin appeared, the mood at a St Petersburg meeting on the future of Russian science was tense. But when Andrei Starinets, an expatriate theoretical physicist now at the University of Oxford, UK, used the former dictator's image to reinforce a call for Russia to lead the way in science — and to ask his fellow émigrés to stand united in “turbulent political times” — tempers exploded.

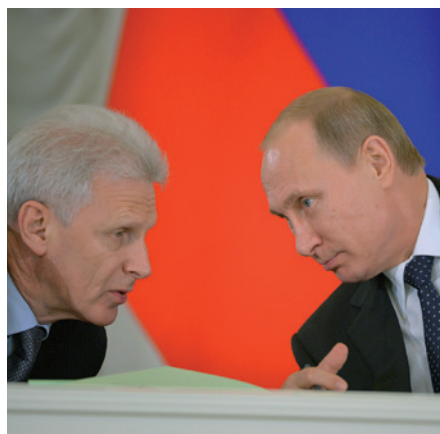
“I’m not going to take this anymore,” shouted Alexey Kondrashov, an expatriate geneticist at the University of Michigan in Ann Arbor. Seething with rage, he jostled his way out of the room and slammed the door behind him.

After Russia's annexation of the Crimea earlier this year and the violent separatism that threatens to tear apart the rest of Ukraine, Starinets' reference to Stalin — whose actions led to the exile and deaths of millions in the gulag system in the first half of the twentieth century — was provocative. But heated debate bubbled up frequently at the meeting, which was convened on 5–6 December by the private European University at Saint Petersburg.

The geopolitical situation has not yet severely hurt collaborations such as the International

Space Station or the ITER fusion reactor that is being built in France. But the gathering, which brought together 100 or so expatriate and resident Russian scientists, as well as government officials, revealed deep divides.

There are those like Starinets, who are staunchly loyal to Russian President Vladimir Putin and think that Russian science can restore its strengths by going it alone. And there are those like Kondrashov, who are deeply worried that their country's recent actions in Ukraine and its weak democracy at



Adviser Andrei Fursenko with Vladimir Putin.

home are making Russia an unpleasant place to do science and are driving away scientists, whether they come from Russia or elsewhere.

“Any discussion about the future of Russian science is pretty much pointless when this country behaves like a bull,” said Kondrashov shortly after he stormed out of the meeting. “I love Russia, but the outlook for science here is gloomy and I’m very concerned about where that country is going.”

One goal of the meeting was to devise ways to restore Russian science. The Soviet Union was a scientific powerhouse, but Russian science is still struggling to recover from its near-collapse in the 1990s, and its output lags behind that of rivals such as China (see ‘Widening gap’). Although Russia has kept its strengths in mathematics and some areas of physics, it trails other large nations in the life sciences.

Tensions at the St Petersburg meeting ran high from the start. On day one, scientists lobbed complaints at Andrei Fursenko, a leading science adviser to the president and one of several close Putin allies on whom the US government imposed sanctions in the spring in response to Russia's actions in Ukraine.

“Do you have a vision for the future of science in this country?” shouted one researcher at Fursenko. “Will we have a say?” another demanded. In part, they were referring to a leaked letter sent by Fursenko to Putin in June, which proposed areas of research to be prioritized for science — and bearing a handwritten “I agree”, apparently from Putin. Many scientists saw that letter as a sign that science policy is being decided behind closed doors, without researchers being consulted.

Meeting attendees also complained to Fursenko about a 2013 reform that put the Russian Academy of Sciences (RAS) under the control of a federal agency that reports directly to Putin. “We are always open about our vision,” replied Fursenko. He added that Russian scientists have a negative attitude towards their government, and he made a promise to increase support for “the best Russian labs” — a sentiment that earned some applause.

Another goal of the meeting was to formulate plans to stem an alarming brain drain from Russia. “Students and intellectuals are fleeing this country,” said Mikhail Gelfand, deputy director of the RAS Institute for Information Transmission Problems in Moscow. One of those was Sergei Guriev, a prominent economist who was a speaker at a similar meeting held four years ago, and who fled Russia over fears about government repressions.

The loss of scientists is not new for Russia: over the past quarter of a century, an estimated 30,000 have migrated to the West and only a few hundred have returned. But many think that the government's current stance is making things worse.

And although there have been positive

ALEXSEY NICHUKHIN/RIA NOVOSTI

ALEXEI DRUZHNIN/RIA NOVOSTI

consequences of this migration for Russia — mainly the links that grow up between Russian and Western labs as a result — there are signs that these links are under strain. “Most of us grew up, studied, and launched our careers in Russia and later benefited from support and political stability in the West,” says Valery Yakubovich, a sociologist and management scholar at the ESSEC business school in Cergy-Pontoise, France. “Maintaining connections is getting more difficult, but even more important in these turbulent times.”

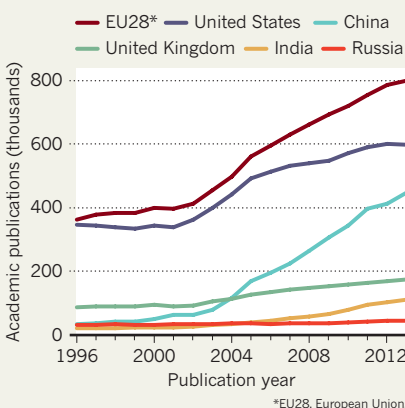
At the meeting there were suggestions that the political climate in Russia is interfering with attempts to lure foreign scientists to work there, and to encourage expatriates to return. In 2010, the government launched a ‘mega-grant’ programme worth 12 billion roubles (US\$428 million at the time) to attract scientists from abroad to do research at Russian universities.

But “why would anyone who lives a decent life abroad decide to do science in Russia at a time when fear and intimidation interfere with everything in this country?” asked Maxim Frank-Kamenetskii, a biomedical engineer at the University of Boston in Massachusetts. He fears that Russia risks falling back to Soviet-era scientific isolation.

Some say that the way to reverse the brain drain is to change things from within. Gelfand

WIDENING GAP

Russia's scientific output trails behind that of other nations and has hardly risen since 1996.



has previously joined Moscow rallies of young Russian scientists and members of the RAS. He called on scientists to have the “moral courage” to create a political environment in which science can flourish. “With a more pronounced civic stance, many bad things here might not happen,” he told the meeting.

But not everyone there saw discussion of politics as fruitful. Elena Grigorenko, an epidemiologist at the Yale Child Study Center in New Haven, Connecticut, and one of only four

women who succeeded in the competition for the mega-grants, chose not to discuss politics. “I’m a Russian citizen and I do care about politics, but it’s my choice when to express my opinions,” she said.

And at least one émigré sees the political situation in Russia as a reason to return. Artem Oganov, a Moscow-born computational materials designer formerly at the State University of New York at Stony Brook, relocated to the Russian capital this month. He will take up a faculty position at the Skolkovo Institute of Science and Technology (Skoltech), an English-language research university that was set up in 2011 in partnership with the Massachusetts Institute of Technology in Cambridge. Oganov is keen to help Russia to restore its science output.

“I’m not a refugee, nobody treated me badly, and I am perfectly at peace with my country,” he says. “I do worry about the sanctions and the growing economic problems here, but I could never forgive myself if Russia needed me and I was not there.” ■

CORRECTION

The News Feature ‘Ebola’s lost ward’ (*Nature* **513**, 474–477; 2014) incorrectly stated that nurse Veronica Koroma contracted Ebola.

2014 IN REVIEW

This year may be best remembered for how quickly scientific triumph morphed into disappointment, and even tragedy: breakthroughs in stem-cell research and cosmology were quickly discredited; commercial spaceflight faced major setbacks. Yet landing a probe on a comet, tracing humanity's origins and a concerted push to understand the brain provided reasons to celebrate.

SPACE RACE EXPANDS Asian nations soared into space this year. The Indian Space Research Organisation put a mission into orbit around Mars — the first agency to do so on its first try. Japan launched the Hayabusa-2 probe, its second robotic voyage to bring back samples from an asteroid. And even as China's lunar rover Yutu (or Jade Rabbit) stopped gathering data on the Moon's surface, mission controllers took the next step in the country's lunar exploration programme by sending a test probe around the Moon and back to Earth.

But for commercial spaceflight, it was a bad year. Virgin Galactic's proposed tourism vehicle *SpaceShipTwo* disintegrated during a test flight in California and killed one of its pilots. That came just three days after a launch-pad explosion in Virginia destroyed an uncrewed

private rocket intended to take supplies to the International Space Station. The accident wiped out a number of research experiments destined for the station, whose managers are trying to step up its scientific output. Problems on the station also delayed the deployment of a flock of tiny Earth-watching satellites, nicknamed Doves, which are part of the general trend of using miniature 'CubeSats' to collect space data.

On a bigger scale, the European Space Agency successfully launched the first in its long-awaited series of Sentinel Earth-observing satellites.

COMETS CALL After a decade-long trip, the European Space Agency's Rosetta spacecraft arrived at comet 67P/Churyumov-Ger-

asimenko in August and settled into orbit. Three months later, Rosetta dropped the Philae probe to 67P's surface, in the first-ever landing on a comet. Philae relayed science data for 64 hours before losing power in its shadowy, rocky landing site.

Meanwhile, a flotilla of Mars spacecraft — probes from India, the United States and Europe — had an unplanned close brush with comet Siding Spring, which zipped past the red planet in October at a distance of 139,500 kilometres — about one-third of the distance from Earth to the Moon. NASA rovers continued to trundle along on the Martian surface: Curiosity finally reached

The BICEP2 telescope at the South Pole may have spied gravitational waves — or dust.

KEITH VANDERLINDE/NSF

the mountain that it has been heading towards since landing in 2012, and Opportunity passed 40 kilometres on its odometer, breaking a Soviet lunar rover's distance record for off-Earth driving.

The search for planets beyond the Solar System also got a huge boost. In February, the team behind the now mostly defunct Kepler spacecraft announced that it had confirmed the existence of 715 extrasolar planets, the largest-ever single haul. Kepler data also revealed the first known Earth-sized exoplanet in the habitable zone of its star, a step closer to the long-sought 'Earth twin'.

HUMAN ORIGINS DECODED Considering that they have been dead for around 30,000 years, Neanderthals had a hell of a year. Their DNA survives in non-African human genomes, thanks to ancient interbreeding, and two teams this year catalogued humans' Neanderthal heritage. Scientists learnt more about the sexual encounters between *Homo neanderthalensis* and early humans after analysing the two oldest *Homo sapiens* genomes on record — from men who lived in southwest Siberia 45,000 years ago and in western Russia more than 36,000 years ago, respectively. The DNA revealed hitherto-unknown human groups and more precise dates for when *H. sapiens* coupled with Neanderthals, which probably occurred in the Middle East between 50,000 and 60,000 years ago. Radiocarbon dating of dozens of archaeological sites in Europe, meanwhile, showed that humans and Neanderthals coexisted there for much longer than was once thought — up to several thousand years in some places.

Genomes old and new charted the emergence of agriculture. Contemporary Europeans carry DNA inherited from light-skinned, brown-eyed farmers who migrated from the Middle East beginning 7,000–8,000 years ago, in addition to more-ancient ancestry. The achievements of these early farmers — domestication of crops such as wheat and barley — are also being understood through genome sequencing. In July, a consortium reported a draft copy of the gargantuan wheat genome, which contains 124,000 genes and 17 billion nucleotides. Another group released the genomes of 3,000 rice varieties.

Genomes of the future may soon carry added information. Scientists in California engineered *Escherichia coli* bacteria to include two chemical nucleotides in their genome in addition to the four that all other life forms use. The next step is to harness the expanded genetic alphabet to produce new kinds of protein. An effort to synthesize an entire yeast genome produced its first chromosome this year.

EBOLA TOLL RISES The Ebola epidemic that ravaged West Africa this year is the largest since the virus was discovered in 1976 — and it exposed major gaps in the world's ability to respond to emerging infectious diseases. By mid-December, around 6,800 people had died in Guinea, Liberia and Sierra Leone.

The first case in the epidemic is thought to be that of a two-year-old in Guinea, who died in early December 2013. A genetic analysis of viral samples suggests that the epidemic began with a single animal-to-human transmission.

Early on, much media attention was focused on experimental drugs, including the antibody cocktail ZMapp, but infectious-disease experts have emphasized the need to expand access to treatment and to implement basic epidemiological measures such as tracing contacts of infected people.

Fears that the epidemic would expand to other countries have proved unfounded; small numbers of cases in Mali, Nigeria, Senegal, Spain and the United States were isolated quickly, and onward spread was limited.

November brought encouraging results from the first safety trials of an experimental Ebola vaccine in healthy human volunteers, with efficacy trials of this and other vaccines set to begin in West Africa early in 2015. Experimental drugs as well as treatments that involve dosing Ebola patients with

'convalescent' blood and serum from survivors are also in tests. But major questions about the virus's biology remain to be answered.

BIG DUST BUST The BICEP2 experiment flexed its muscles in March, when astronomers reported evidence of gravitational waves from the Big Bang — seeming confirmation of cosmic inflation, the initial exponential expansion of the Universe. But it quickly emerged that the BICEP2 radio telescope, located at the South Pole, may actually have detected a signal distorted by cosmic dust; this theory is supported by results from the European Space Agency's Planck satellite, announced in September. The BICEP2 and Planck teams are set to release a joint analysis soon that should provide a definitive answer to the gravitational-wave quandary.

China advanced plans for an electron-positron supercollider to study the Higgs boson, and is considering an even more ambitious goal: a next-generation super proton-proton collider at the same, as-yet-unbuilt facility.

Graphene showed new-found vulnerability, as scientists discovered that the material — the world's thinnest and strongest — allows protons to pass through it. This suggests new applications in hydrogen fuel cells, or perhaps a membrane that can collect hydrogen from air. (On a lighter note, *Nature Materials* published a recipe in April for how to make ►



At the one-year mark, there is no telling when the Ebola epidemic in West Africa will end.

THE YEAR IN SCIENCE: WINS AND LOSSES



30 January

Birth of first genetically engineered monkeys is announced.

WINS



23 March

Water is released into the Colorado River delta for the first time in decades, as part of an ecological experiment.

7 May

First living cell created with synthetic DNA building blocks in its genome.

LOSSES

2 January

"LET'S BE CLEAR. US BECOMING LOCKED IN ICE WAS NOT CAUSED BY CLIMATE CHANGE."

Chris Turney, leader of a private polar expedition aboard the *Akademik Shokalskiy*, which was stranded in Antarctic ice for a week.

24 February

Publishers announce intent to withdraw more than **120** papers after discovering documents were computer-generated nonsense.

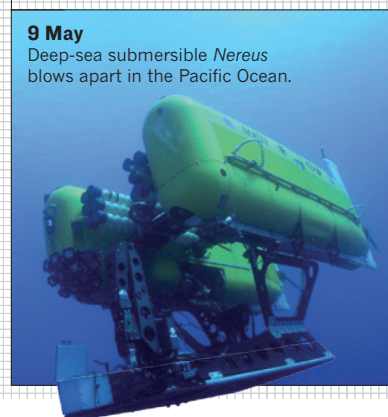
30 April

World Health Organization warns that world may be heading into a "post-antibiotic era".



9 May

Deep-sea submersible *Nereus* blows apart in the Pacific Ocean.



► graphene in a kitchen blender — but for various disappointing yet practical reasons, the recipe is not recommended for home use.)

HIV HOPES DASHED For HIV researchers, 2014 brought a steady stream of bad news. Last year, physicians said that the 'Mississippi baby', a child born with HIV, was cured by aggressive early treatment with antiretroviral drugs. But in July, researchers announced that the child, now four years old, had detectable levels of HIV in her blood. Her story echoed that of two men treated in Boston, Massachusetts, who had been virus-free for several years after bone-marrow transplants; in December 2013, word came that they had relapsed.

In July, the International AIDS Conference in Melbourne, Australia, was rocked by the loss of six delegates, including famed clinical virologist Joep Lange of the University of Amsterdam. They died en route to the meeting when their plane — Malaysia Airlines flight MH17 — was shot down over Ukraine.

A few promising results emerged this year, such as the debut of a treatment that makes immune cells resistant to HIV by editing their DNA, and the discovery of two HIV-infected Australian men whose stem-cell treatments for cancer rendered their virus undetectable — so far.

BRAIN GAINS Unprecedented advances in nanotechnology and computing have helped to drive the emergence of ambitious projects to understand the brain. This year, many such efforts reached major turning points — not all of them positive. In July, the European Union's flagship project to model the brain in a supercomputer faced a mutiny. In a protest letter to the European Commission, more than 150 key scientists charged that the billion-euro Human Brain Project had become autocratic and was



A dry California waited for an El Niño to bring rain.

veering away from its scientific goals; they threatened to withdraw their cooperation unless the programme's management was overhauled. The sides are now in mediation, and a revised research plan is expected early in 2015.

More peaceably, the US BRAIN (Brain Research through Advancing Innovative Neurotechnologies) Initiative awarded its first grants this year. Japan joined the global brain-wave in October, when it announced a bold ten-year project called Brain/MINDS (Brain Mapping by Integrated Neurotechnologies for Disease Studies) that will map the marmoset brain to aid studies on human neurological and psychiatric diseases.

MERCURY RISING For many climate scientists, the past several months have been a frustrating exercise in waiting for an El Niño — a powerful warming event in the eastern Pacific that was forecasted, but which never arrived. Even so, 2014 is likely to rank as the hottest since modern records began about 140 years ago — just beating 1998, 2005 and 2010, which are in a statistical dead heat.

Scientists still debate the causes of the relatively slow warming trend over the past 15 years. One thought-provoking study published this year attributes the warming pause to periodic changes in ocean circulation that

MONKEYS: NIU ET AL.; CELL: RIVER; PETE MCBRIDE/NATL GEOGRAPHIC SOC./CORBIS; PLUS: ANDREW AITCHISON/CORBIS; SUBMERSIBLE: ANL/WHOI

DAVID MCNEW/GETTY



12 August

Maryam Mirzakhani becomes first woman to win a Fields Medal (pictured).

11 September

"WE HAVE FINALLY ARRIVED."

John Grotzinger, leader of NASA's Curiosity rover programme, on news that the rover had reached Mount Sharp after more than two years of driving across Mars.

10 November

Six seismologists cleared of manslaughter in charges stemming from a 2009 earthquake in L'Aquila, Italy.

2 October

World's first commercial coal-fired power plant that can capture its carbon-dioxide emissions is officially launched in Canada.

12 November

European Space Agency's Rosetta mission makes history by landing a probe on a comet.

17 July

Malaysia Airlines flight MH17 is shot down in eastern Ukraine, killing six people on their way to an international AIDS conference.

10 July

"IT FELT VERY MUCH LIKE A PUNCH TO THE GUT."

Hannah Gay, a paediatric HIV specialist, on news that the 'Mississippi baby' — thought to have been cured of HIV by drug treatment — had relapsed.

29 August

Icelandic volcano erupts. By mid-October, Bárðarbunga was spewing twice as much sulphur dioxide each day as all of Europe's smokestacks.



3 December

World Meteorological Organization says that 2014 is on track to become the warmest year on record.

MEDAL: STEFAN ZACHOW; VOLCANO: LUKAS GAWENDA/IMAGEBROKER/CORBIS

carry heat to the deeper layers of the Atlantic and Southern oceans. Another analysis argues that the slowdown has been driven by warming of the Atlantic Ocean, which then caused the eastern Pacific to cool.

On the policy front, the Intergovernmental Panel on Climate Change wrapped up its fifth assessment report in November, warning of "severe, pervasive and irreversible impacts for people and ecosystems" if greenhouse-gas emissions continue. The United States and China seemed to declare a climate truce, with fresh pledges to reduce their greenhouse-gas emissions — raising hopes that developed and developing nations will meet their goal of agreeing on a new international climate treaty at talks in Paris in 2015.

STEM-CELL DRAMA The year started out with a stem-cell boom. In January, researchers at the RIKEN Center for Developmental Biology (CDB) in Kobe, Japan, announced a surprising discovery: an unexpectedly fast and easy way to make pluripotent stem cells by immersing mature cells in acid or applying physical pressure. But the studies, published in *Nature*, were found to contain manipulated

figures and images, and efforts to replicate them failed. The papers were retracted in July. In August, a co-author from the CDB — Yoshiki Sasai, a pioneer of regenerative medicine — took his own life.

In September, the beleaguered CDB got a bit of good news, as centre ophthalmologist Masayo Takahashi led the first clinical trial of induced pluripotent stem (iPS) cells. Fresh hope was also given to the world's first clinical trial of embryonic stem cells to treat spinal-cord injury, which was restarted after an abrupt shutdown in 2011.

In another advance, Douglas Melton of Harvard University in Cambridge, Massachusetts, worked out how to create insulin-creating β -cells from stem cells — a finding that could lead to new treatments for type 1 diabetes if researchers can keep the immune system from attacking the cells. Meanwhile, 'right-to-try' laws have emerged in several US states, allowing the use of unproven stem-cell therapies; in Japan, new clinical guidelines allow stem-cell treatments to enter the clinic without a rigorous efficacy trial — raising bioethical concerns.

FRIGHTENING FINDS What biohazards live in your refrigerator? On 1 July, US government researchers found six vials of 60-year-old smallpox virus in a storage room

at the US National Institutes of Health (NIH) campus in Bethesda, Maryland.

The discovery drew attention to biosafety lapses at US government laboratories, including revelations by the Centers for Disease Control and Prevention that its researchers had mishandled anthrax spores and accidentally shipped dangerous H5N1 influenza virus to another laboratory. In August, an NIH 'safety sweep' produced a 100-year-old box that contained dangerous pathogens and the toxin ricin.

The incidents renewed debate on whether the benefits of some pathogen research are outweighed by its potential dangers. In mid-October, the White House shocked researchers by announcing that it would not fund new 'gain-of-function' studies that engineer pathogens — such as influenza virus — to become more deadly or transmissible. It also asked researchers to pause ongoing gain-of-function experiments. The NIH went a step further — it ordered about 20 projects that it funds to halt while two advisory groups examine the risks and benefits of such research over the next year. ■

Written by Lauren Morello, Alison Abbott, Declan Butler, Ewen Callaway, David Cyranoski, Sara Reardon, Quirin Schiermeier & Alexandra Witze

365 DAYS:
the year in science

IMAGES OF THE YEAR

Incredible discoveries in 2014 arose from researchers' relentless pursuit of answers about the world. From the far reaches of space to the depths of the oceans, *Nature's* selection of this year's most striking images document both natural disasters and technological wonders.

Images selected by *Nature's* art and design team
Text by Daniel Cressey



VOLCANO TRAGEDY

Mount Ontake, an active volcano some 200 kilometres west of Tokyo, has long been a popular tourist destination in Japan. Despite careful monitoring by scientists, an eruption on 27 September caught many off guard, spraying ash and debris over the surrounding region and killing more than 50 people. Rescue teams battled thick ash to search for survivors in remote lodges near the mountain's peak.

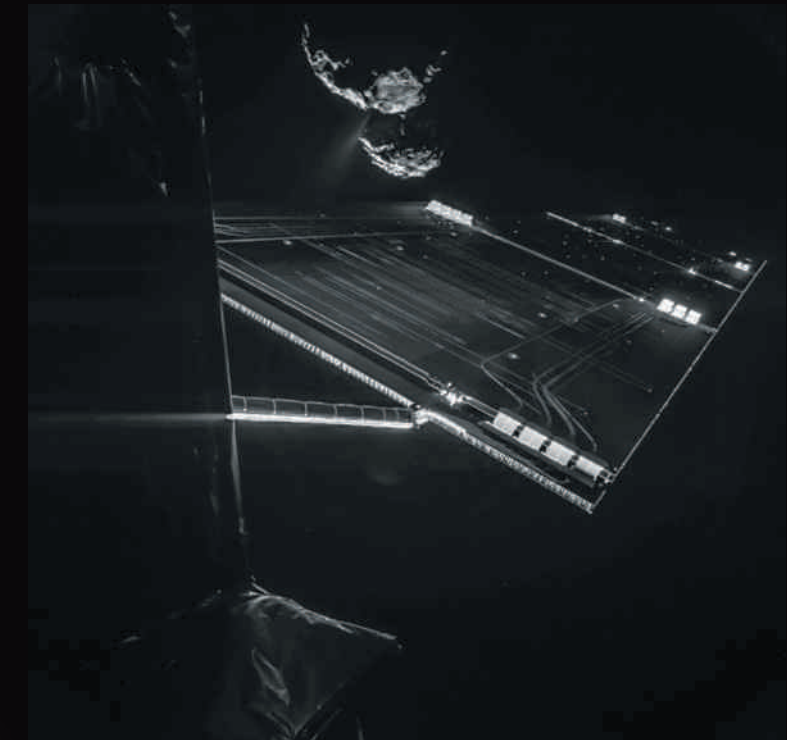
MOON MOSAIC

Jupiter's moon Europa, as it would look to human eyes. NASA reprocessed a series of images taken by the Galileo space probe in the late 1990s, adjusting the colours to create this realistic, high-resolution view of the moon's icy terrain.



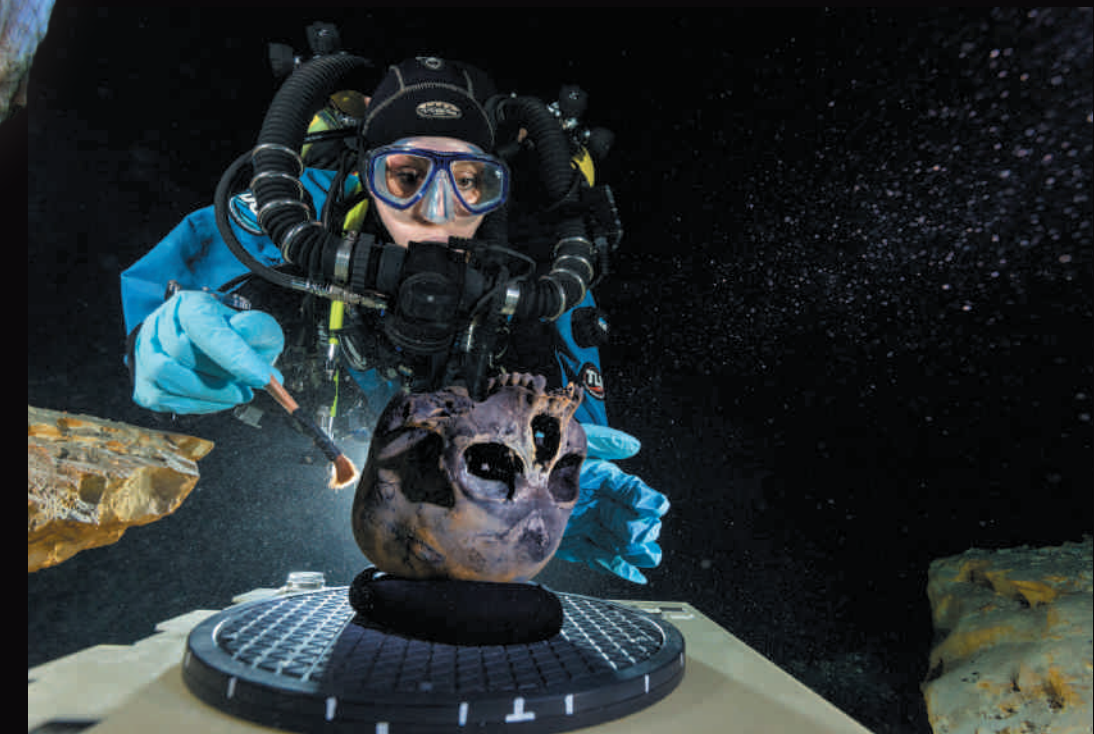
DUMBO AT DEPTH

Tentacles coiled in a pose never seen before, this 'dumbo octopus' of the genus *Grimpot euthis* was captured on camera in April in the Gulf of Mexico. Researchers on the US vessel *Okeanos Explorer* got this rare glimpse of the creature by piloting a remote-controlled submersible to a depth of some 2,000 metres.



PHILAE IN FLIGHT

The world was on tenterhooks in November as the European Space Agency's Rosetta spacecraft attempted to put the Philae lander on the surface of comet 67P/Churyumov-Gerasimenko. Before successfully completing the tricky manoeuvre, Philae sent back this picture of itself closing in on its target as they both moved through space at more than 50,000 kilometres per hour.



EXTREME ARCHAEOLOGY

The 12,000-year-old skull of a teenager from Mexico sits on a rotating platform, enabling divers to take a three-dimensional scan of the remains. Found deep inside submerged caves in Mexico's Yucatán Peninsula, the skull is part of a remarkable collection of ancient bones that are helping to shed light on how humans spread across the Americas. Difficulties in removing the remains meant that divers had to analyse them *in situ*.



A FURIOUS FISSURE

Dawn and dusk in Iceland turned blood red earlier this year as volcanic pollution filled the skies. The Holuhraun fissure — near the erupting Bárðarbunga volcano — belched out thousands of tonnes of sulphur dioxide every day, surprising scientists who were expecting ashy expulsions similar to the Eyjafjallajökull eruption in 2010.



EYE TO EYE

These staring eyes of a *Phidippus audax* jumping spider secured third place in Nikon's Small World photography competition for Noah Fram-Schwartz of Greenwich, Connecticut.

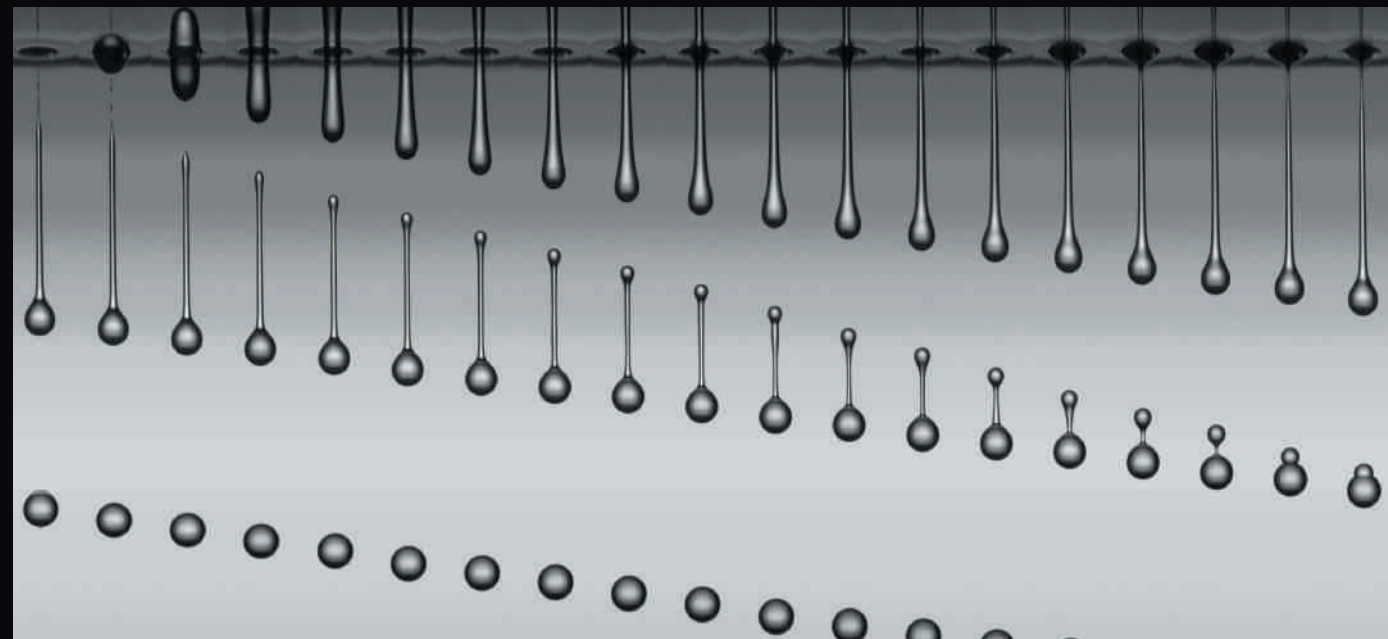
THE SEE-THROUGH MOUSE

Why scan bodies if you can just make tissue transparent? This mouse has been rendered see-through by a team in Japan using a chemical cocktail and computational imaging, one of a growing number of methods that reveal organs without dissection.



ABOVE THE CLOUDS

It remains unclear why Japanese artist Azuma Makoto attached a bonsai tree to a balloon and launched it into the upper layers of the atmosphere. But the result of his 'Exobotanica' project, which has now sent numerous plants into space, was a series of beautiful pictures like this one.



DROP EVERYTHING

These picolitre-sized silicone-oil droplets were snapped by researchers at the University of Twente in the Netherlands. Lighting up the droplets with 8-nanosecond-long laser pulses, the team took images 600 nanoseconds apart to capture how the falling droplets formed.



SHINE ON

The eerie green glow in this image from Emas National Park in Brazil emanates from the bioluminescence of click-beetle larvae living on a termite mound — and from the flight paths of adult beetles. Photographer Ary Bassous's long exposure made him the winner in the invertebrates category of this year's Wildlife Photographer of the Year competition, run by London's Natural History Museum and BBC Worldwide.

HOLUHLRAUN: RAGNAR TH. SIGURDSSON/ARCTIC IMAGES.COM; MOUNT ONTRAKE: KYODO NEWS/AP; EUROPA: NASA/JPL-CALTECH/SETI INST.; DUMBO: NOAA/CC BY SA 2.0; PHILAE: ESA/ROSETTA/PHILAE/CIVA; SKULL: PAUL NICKLEN/WATL GEOGRAPHIC; SPIDER: NOAH FRAM-SCHWARTZ/ COURTESY OF NIKON SMALL WORLD; SEE-THROUGH MOUSE: COURTESY OF RIKEN; BONSAI: AMKK/EXOBOTANICA; DROPLETS: MARK J. VAN DER MEULEN & ARJAN VAN DER BOSS; BEETLE LARVAE: ARY BASSOUS/WILDLIFE PHOTOGRAPHER OF THE YEAR 2014.

NATURE'S 10

Ten people who mattered this year.



CGI ILLUSTRATION BY PETER CROWTHER ASSOCIATES C/O DÉBUT ART

ANDREA ACCOMAZZO / SUZANNE TOPALIAN / RADHIKA NAGPAL /
MASAYO TAKAHASHI / DAVID SPERGEL / MARYAM MIRZAKHANI / PETE FRATES /
KOPPILLIL RADHAKRISHNAN / SHEIK HUMARR KHAN / SJORS SCHERES

365 DAYS:
the year in science

COMET CHASER

A former test pilot steered the Rosetta mission to an icy world in deep space.

ANDREA ACCOMAZZO

BY ELIZABETH GIBNEY

Nearly two decades ago, Andrea Accomazzo got into trouble with his girlfriend when she found a scrap of paper on his desk. In his handwriting was scrawled a phone number next to a female name: Rosetta.

"She thought it was a girl," says Accomazzo. "I had to explain to my jealous Italian girlfriend that Rosetta is an interplanetary mission that is flying to a comet in almost 20 years."

Ever since, Accomazzo has divided his attention. He eventually married his girlfriend and has also spent the past 18 years pursuing the comet 67P/Churyumov-Gerasimenko. As flight director for the mission, Accomazzo led the team that steered Rosetta to its August rendezvous with the comet, following a 6.4-billion-kilometre journey from Earth. The pinnacle of the project came in November, when Rosetta successfully set down a lander named Philae, providing scientists with the first data from the surface of a comet and making it one of the most successful missions in the history of the European Space Agency (ESA).

Accomazzo did not act alone: it took a large operations team at ESA to manoeuvre Rosetta with enough precision to drop Philae down just 120 metres from the centre of the landing zone. "Given that we'd had a 500-metre error circle, that was not a bad shot," says Fred Jansen, who led the mission. When Philae's anchoring systems failed, the craft bounced into a shady site where it could not charge its solar panels, so the lander lost power after 64 hours. But in that time, it gathered a trove of data that will add to the information collected by Rosetta about the comet's structure and composition. Armed with those insights, scientists hope to better understand the origin and evolution of the Solar System, including whether comets could have brought water and

organic molecules to Earth during its infancy.

Accomazzo started off his career focused on a different type of flight. He first trained as a test pilot in the Italian Air Force. But although he loved flying, he found the culture too constraining and after two years he quit to study aerospace engineering. With his quiet, hard-working, sometimes no-nonsense nature, colleagues say that Accomazzo brings a bit of the military with him into mission control.

For Accomazzo, the biggest parallel between flying a fighter jet and Rosetta is the need for split-second judgements. "You have to prepare and train a lot to be able to make the right decision, very quickly," he says. Between launch and landing, his team ran 87 full-day simulations.

Although the Rosetta mission has been a broad success, Accomazzo still cried when he heard that Philae had died, and hopes the lander will revive when the comet approaches the Sun. After swinging around the Sun in August 2015, the comet will head back out towards deep space.

By early 2017, there will be too little sunlight to power Rosetta, and Accomazzo is planning a daring finale. He would love to see the craft skim above the surface of the rubber-duck-shaped comet through the valley that separates its body and head. The team might even try to land the spacecraft on the comet's surface.

The decision might not be up to him. Accomazzo is stepping away from the day-to-day flight operations at Rosetta and is busy preparing for ESA's interplanetary missions to Mercury, Mars and Jupiter. Even with such exciting projects, he finds it hard to leave Rosetta behind. "It's a bit sad," he says. "I don't know how I will be able to cope."

He still dreams of Rosetta. "This morning I woke up at 4 a.m. and thought 'something is wrong,'" he says. "At 7.30 a.m. I got a call — Rosetta had briefly lost signal to Earth at 4 a.m. — I often have this kind of episode. I'm totally linked." ■

ANDREA REEG/AGENCY FOCUS/EYEVINE

CANCER COMBATANT

One clinician always believed that cancer immunotherapy would work — and she was right.

BY HEIDI LEDFORD

When Suzanne Topalian heard in July that a therapy she had helped to pioneer could now be used in the United States to treat people with advanced melanoma, she greeted the news with excitement, but also characteristic resolve. The meticulous cancer researcher and physician was already focused on the field's next challenges: approval for the drug in other countries and against a wider range of cancers. "Although this was reason to celebrate, we're still looking towards the horizon," she says.

The drug in question is part of a hot new class called PD-1 inhibitors, which allow T cells in the immune system to jump into high gear so that they are free to attack tumours. This July, Japanese regulators approved the first such drug — nivolumab, made by Bristol-Myers Squibb of New York — largely on the back of clinical trials that Topalian led. Two months later, the US Food and Drug Administration approved another, called pembrolizumab. Some

analysts predict that the drugs will become a cornerstone of cancer treatment, with a market exceeding US\$10 billion by 2020.

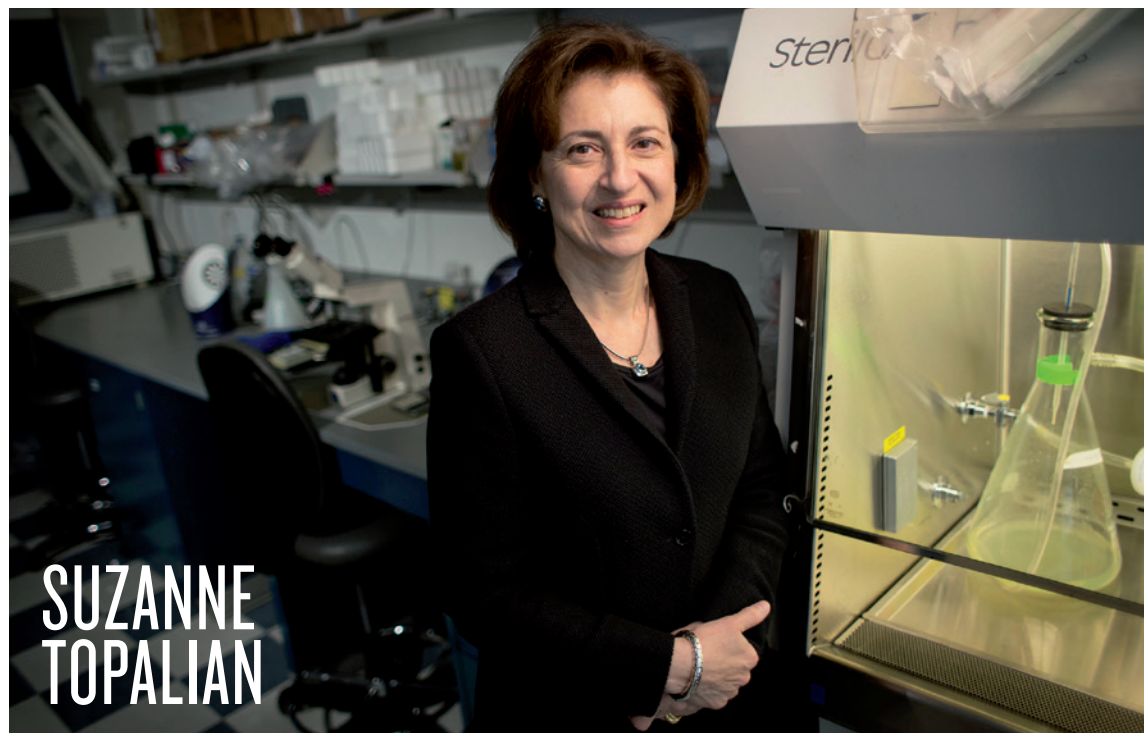
Even as a medical student, Topalian says, she was hooked by the idea of turning the body's own defences against cancer rather than — as most other therapies do — attacking a tumour directly with radiation or drugs. In 1985, she joined the lab of tumour immunologist Steven Rosenberg at the US National Cancer Institute in Bethesda, Maryland. She intended to leave after 2 years; instead, she stayed for 21 and set up her own lab. Rosenberg says that Topalian quickly made her mark as a talented, careful scientist who always kept the big picture in mind. "She was totally passionate about finding effective cancer treatments," he says.

Even when sceptics doubted that cancer immunotherapy would work, and early clinical trials looked disappointing, Topalian was undeterred. "There would always be some patients who responded to those treatments," she says. "It was those exceptional responders who kept hope alive."

In 2006, Topalian left Bethesda to help to launch trials of nivolumab at Johns Hopkins University in Baltimore, Maryland. That work led to a landmark publication in 2012 showing that nivolumab produced dramatic responses not only in some people with advanced melanoma but also in those with lung cancer — the world's most common cause of cancer death (S. L. Topalian *et al.* *N. Engl. J. Med.* **366**, 2443–2454; 2012). Regulators are now considering approval of the drugs for treatment of lung cancer.

Other researchers are pouring into the field, spurred by successes with PD-1 inhibitors and other cancer immunotherapies, says Jedd Wolchok, an oncologist at the Memorial Sloan Kettering Cancer Center in New York. "It's legitimized a field that was once scorned," he says. ■

CHRIS MADDALONI



SUZANNE
TOPALIAN



**RADHIKA
NAGPAL**

ROBOT-MAKER

A researcher inspired by social insects gets robots to coordinate on a massive scale.

BY CORIE LOK

When Radhika Nagpal was a high-school student in India, she hated biology: it was the subject that girls were supposed to study so that they could become doctors. Never being one to follow tradition, Nagpal was determined to become an engineer.

Now she is — leading an engineering research team at Harvard University in Cambridge, Massachusetts. But she also has a new appreciation for the subject she once disliked. This year, her group garnered great acclaim for passing a milestone in biology-inspired robotics.

Taking their cue from the way in which ants, bees and termites build complex nests and other structures with no central direction, Nagpal's group devised a swarm of 1,024 very simple 'Kilobots'. Each Kilobot was just a few centimetres wide and tall, moved by shuffling about on three spindly legs and communicated with its immediate neighbours using infrared light. But the team showed that when the Kilobots worked together, they could organize themselves into stars and other two-dimensional shapes (M. Rubenstein *et al. Science* **345**, 795–799; 2014). Achieving that level of cooperation in a swarm this large was a major feat, says Alcherio Martinoli, a roboticist at the Swiss Federal Institute of Technology in Lausanne. Nagpal's approach — combining theoretical proofs with a physical demonstration of swarm behaviour — "is, to me, extremely powerful and something other people should follow", he says.

The hope is that this kind of swarm-robotics research will eventually

REFLECTION FILMS

EBOLA DOCTOR

An infectious-disease expert battled a killer virus in Africa.

BY ERIKA CHECK HAYDEN

In this year's devastating outbreak of Ebola, Sheik Humarr Khan played a unique part. He was a scientist — part of the team that performed the first genetic sequencing studies of the virus in his native Sierra Leone. He was an infectious-disease doctor who turned down an invitation to leave his country so that he could stay and treat patients. He also became one of its many victims, dying on 29 July.

Ebola brought devastation to Guinea, Sierra Leone and Liberia as it ballooned into an epidemic during 2014. Khan was the lead physician at Sierra Leone's Kenema Government Hospital, where he was treating and studying Lassa, another potentially fatal viral disease, until the hospital was overwhelmed by people with Ebola.

According to those who knew him, Khan believed that research and medicine should serve everyone — not just those able to access and

afford it — and he had eschewed offers to make more money working in the capital, Freetown, to stay in the underserved rural region of Kenema. "That was one of the more important examples he set," says John Schieffelin, a physician at Tulane University in New Orleans, Louisiana, who worked with Khan.

Khan became a central figure in the Kenema community and when Ebola struck, he cancelled his plans to teach abroad. When he became sick himself, his doctors decided not to give him the experimental treatment known as ZMapp in case it backfired and caused dangerous side effects. Some staff at the hospital worried that his death would spark civil unrest. "They said that if Dr Khan dies, people in Kenema are going to tear the hospital down," remembers Lina Moses, an epidemiologist also at Tulane who spent much of the year working in Kenema.

The outbreak now looks as though it is levelling off, and drug and vaccine trials are getting under way. The research that Khan was involved in showed how quickly the virus was mutating, and the team he worked with is now installing genetic sequencers throughout West Africa so that they can continue to track its evolution.

But the toll has been great: Ebola has killed around 6,300 people, including many doctors and other health-care workers. Recovering from this loss of scarce experts will be a tremendous challenge, says Estrella Lasry, a tropical-medicine specialist for Médecins Sans Frontières (Doctors without Borders) in New York City. "It's going to take years before the same number of people who died are trained". ■

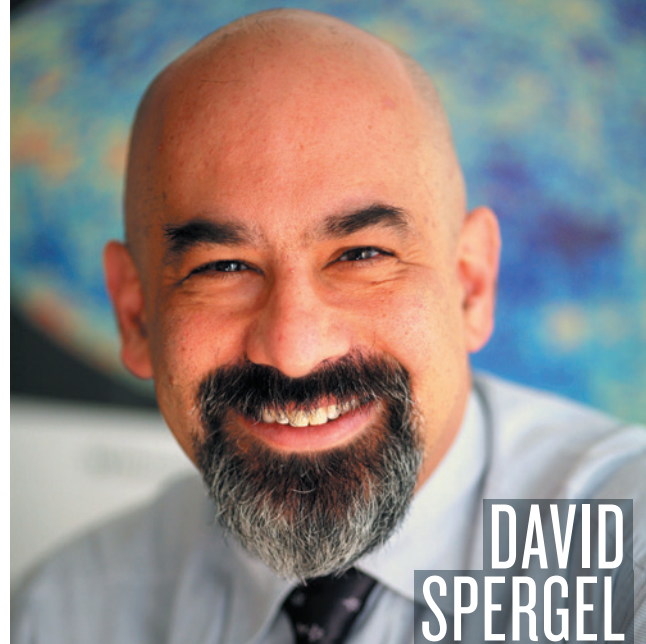
lead to self-organizing robot teams that can rapidly respond to disasters, say, or aid in environmental clean-up. But getting even this far took much longer than Nagpal and her team originally estimated.

The original idea for the Kilobots is four years old, says Nagpal. Like other swarm-robotics researchers, her team had been doing computer simulations and small laboratory experiments. But then one of her postdocs, Michael Rubenstein, convinced her that it was possible to do much larger experiments, because advances in electronics, materials and three-dimensional printing were making it easier and cheaper than ever to create robots en masse.

The team struggled to go from building 20 autonomous robots — their largest group at the time — to the full-sized swarm of 1,024 Kilobots. The key turned out to be simplicity, says Nagpal. “The individuals would be less calibrated, have lower-quality components and would have less control over what they do,” she says, but they would still need to carry out complex tasks by working together. “Somehow, at the top, we would have to think of algorithms that didn’t depend on precision at the individual level.”

Nagpal is now trying to develop large robot swarms that can self-assemble into structures in three dimensions. And she will continue to draw her inspiration from nature, she says — a practice she learned from her graduate-school adviser, computer scientist Gerald Sussman at the Massachusetts Institute of Technology in Cambridge. Sussman convinced her to set aside her distaste for biology when he pointed out that cells are the ultimate computers, able to take in data from signalling molecules, and to carry out complex chemical calculations to decide how to act. And then there are the extraordinary things that happen when these cell-computers come together, says Nagpal.

“At the end, you get this functioning organism and it’s so amazing that you forget that it’s even composed of cells,” she says. This is a key goal in swarm-intelligence research: using the collective to accomplish much more than the individual. “Looking at biology makes me think differently about computer science,” she says. ■



COSMIC SCEPTIC

An astrophysicist found errors in a major discovery about cosmic inflation.

BY RON COWEN

David Spergel first spotted the blunder while on a train in late March. Ten days earlier, researchers had made front-page headlines by holding a press conference announcing the probable detection of gravitational waves from the far reaches of space. That long-sought signal provided evidence that the infant Universe had undergone a brief but enormous expansion called cosmic inflation, and the result had prompted talk of a Nobel prize for the team, which was led by John Kovac of the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts.

Spergel was troubled from the start by the evidence that Kovac’s team had gathered from the BICEP2 telescope at the South Pole. As an astrophysicist who studies the early Universe at Princeton University in New Jersey, he worried that the signal might be an artefact. On the train, en route to giving a lecture in New York City, he realized that the BICEP2 team had made a mistake when accounting for how nearby dust might alter the long-distance signal. He raised his concerns in his talk, and in May he co-authored a paper that pointed out the flaws (R. Flauger *et al.* Preprint at <http://arxiv.org/abs/1405.7351>; 2014).

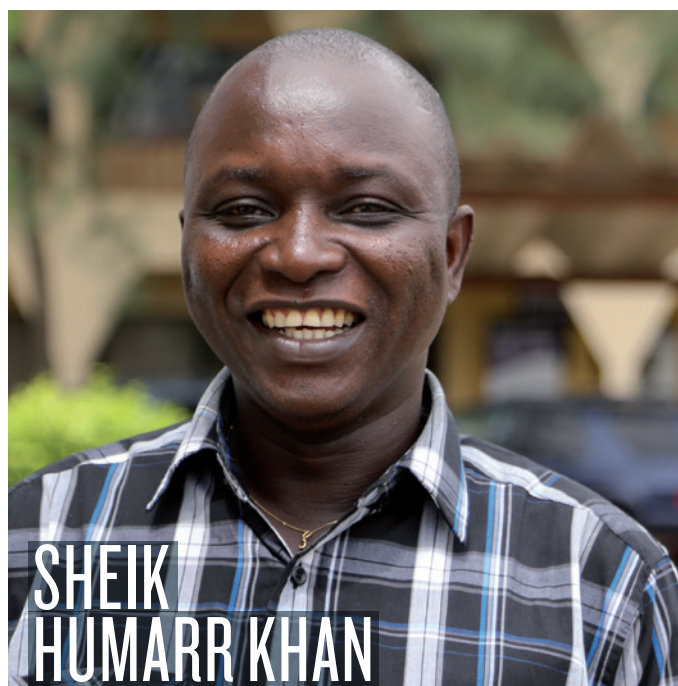
Spergel, who sports a shaved head and a voice that can fill a room, decided that he needed to speak out. “I wanted to let the broader physics community know there were reasons to have doubts,” he says.

Social media amplified his criticisms. A video of his New York talk drew nearly 2,000 views, alerting others to the controversy. Soon, talk of a Nobel prize for the BICEP2 team was eclipsed by discussions about how it had made a cosmic mistake.

When the BICEP2 researchers published their findings in June (P. A. R. Ade *et al.* *Phys. Rev. Lett.* **112**, 241101; 2014), they were more tentative than at the press conference — although not enough to satisfy Spergel. A forthcoming analysis of satellite data may soon settle the controversy. For cosmologist Marc Kamionkowski of Johns Hopkins University in Baltimore, Maryland, the episode shows the danger of announcing major results too early. Although the BICEP2 researchers may have had good reasons to hold a press conference, he says, “they or others in a similar situation in the future may lean towards awaiting some vetting.” ■

JOHN JAMESON/PRINCETON UNIV.

PARDIS SABETI





MARYAM
MIRZAKHANI

SURFACE EXPLORER

A mathematician's award shines a light on a lack of women in the field.

BY ERICA KLARREICH

When Maryam Mirzakhani was a mathematics graduate student at Harvard University in 2003, she went to her adviser, Curtis McMullen, with a question. McMullen had just solved a long-standing problem related to the behaviour of billiard balls on a type of abstract table that can be folded up into a doughnut surface with two holes. It was a major discovery, but Mirzakhani asked why he had proved it just for surfaces with two holes, rather than for complex surfaces with even more. She was drawn to the largest possible problem — even if she had no idea, back then, just how hard it would be to solve. “Maybe sometimes not knowing enough is a blessing,” she says, “because then you just do your thing.”

Mirzakhani, now at Stanford University in California, turned this problem over in her mind for almost a decade, until she found an answer. In a 172-page paper written in 2012 with Alex Eskin of the University of Chicago, Illinois, she extended McMullen's result to all surfaces with two or more doughnut holes, tying together disparate mathematical fields such as geometry, topology and dynamical systems (A. Eskin and M. Mirzakhani Preprint at <http://arxiv.org/abs/1302.3320>; 2013). “It's a spectacular result,” says Howard Masur, a mathematician at the University of Chicago. In August, Mirzakhani was awarded the Fields Medal, often called mathematics' Nobel prize, for this and other advances in pure mathematics. Among her other findings is a surprising link between hyperbolic

COURTESY OF MARYAM MIRZAKHANI

ICE-BUCKET CHALLENGER

A patient advocate helped to kick-start the social-media stunt of the year — with huge returns for research.

BY SARA REARDON

In the two-and-a-half years since he was diagnosed with amyotrophic lateral sclerosis (ALS), 29-year-old Pete Frates has lost the ability to speak or move. But in November, the former university baseball coach was the guest of honour at a sporting-goods shop near his home in Beverly, Massachusetts, where he sat with his newborn daughter in his lap and watched a Christmas celebration that featured an actor dressed as Santa Claus dousing himself with snow.

Santa was paying homage to the ‘Ice Bucket Challenge’, in which people post and share videos of themselves dumping ice water over their heads to raise awareness and donations for ALS research. Frates first promoted the idea in August, through posts to Facebook and YouTube that he dictated using eye-tracking software. Since then, it has become one of the most

lucrative social-media fund-raisers ever for biomedical research — and has led advocates for other little-known diseases to wonder whether similar efforts could also help them to raise money.

The ice-bucket idea did not originate with Frates's posts — similar challenges had been used in other social-media campaigns. But his efforts, along with posts by Pat Quinn of Yonkers, New York, who also has ALS, did a lot to help the challenge go viral. Both men urged Internet users to show solidarity by posting videos. The meme morphed into a fund-raising campaign: either dump water on your head or donate money to ALS research, then challenge friends to do the same. Many people chose both.

So far, participants from around the world have posted at least 17 million ice-bucket videos on Facebook, and raised more than US\$115 million — almost three times the \$40 million the US National Institutes of Health spent on ALS research last year.

Critics say that the Ice Bucket Challenge is a fad and that its focus on a disease affecting around 500,000 people worldwide could draw attention away from deadlier threats, such as heart disease, which kills 7.4 million people every year. Nevertheless, the strategy has caught the attention of other advocacy groups. The National Organization for Rare Disorders in Danbury, Connecticut, held a seminar in October on viral fund-raising campaigns, and is planning a follow-up session owing to its popularity.

Back in Massachusetts, the Frates family still hopes that the Ice Bucket Challenge will one day pay off for ALS. “When there is a treatment,” says Pete's father, John Frates, “it will go back to August 2014.” ■

geometry — the geometry of saddle shapes — and string theory.

Mirzakhani is humble — when she got word of her award, she assumed it came from a hacked e-mail account — and extremely private. She kept a low profile after her prize was announced, but the news was greeted with an explosion of interest elsewhere. It raced through social media and the press, reaching outlets such as the fashion magazine *Elle* and the feminist blog Jezebel. Most of the discussion was not about abstract surfaces, however: it was about how the Iranian-born mathematician was the only woman to receive the Fields Medal since the prize was first awarded in 1936.

The commotion threw a spotlight on the vast under-representation of women in mathematics: according to a 2012 survey of US universities by the American Mathematical Society, women make up only 30% of PhD students — a number that has not budged for years — and only 12% of tenured faculty members at PhD-granting universities. Those who do become tenured mathematics professors receive a disproportionately small number of scholarly awards.

Mirzakhani says that she has not encountered any outright discrimination against women, but that there are subtle cultural forces that can undermine their confidence, such as a shortage of peers and a perception among girls that mathematics isn't "cool". She hopes her award will inspire confidence in female mathematicians — and others believe that it will change how they are perceived. From now on, "no one will be able to think about the Fields Medal without picturing Maryam Mirzakhani", says Ruth Charney, a mathematician at Brandeis University in Waltham, Massachusetts, and president of the Association for Women in Mathematics. "It's a clear signal that there are women doing absolutely top-notch mathematics — in case anyone wasn't sure."

Mirzakhani is sure, and she predicts more female Fields Medal winners soon. Meanwhile, she is focusing on pushing her analysis of billiard surfaces even further. She regards herself as a discoverer, not an inventor, of mathematics. "I see it as exploring some unknown territory," she says. "It's an adventurous thing, trying to find the connections." ■



**KOPPILLIL
RADHAKRISHNAN**

ROCKET LAUNCHER

India's space chief led the country's charge to Mars.

BY T. V. PADMA

Koppillil Radhakrishnan knew the odds were against him when India's Mangalyaan space probe closed in on Mars this year. As head of the Indian Space Research Organisation (ISRO), he was well aware that half of all attempts to reach Mars have ended in failure.

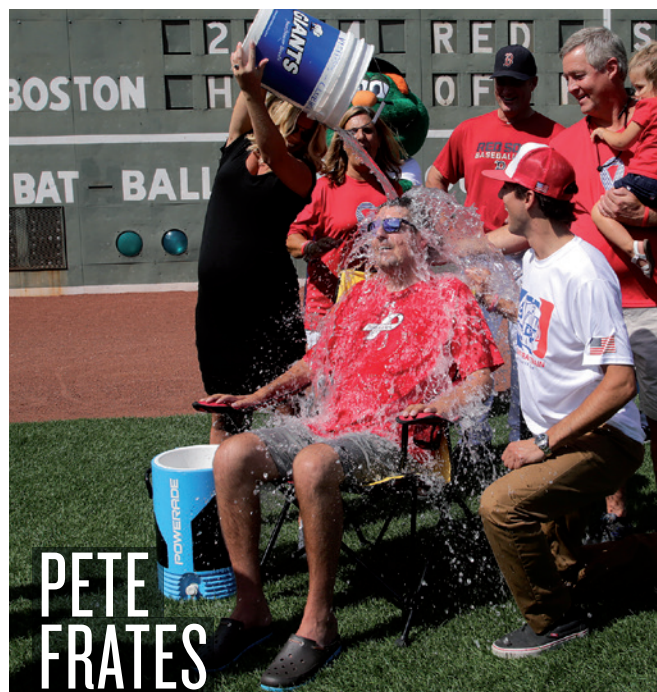
But the ISRO had taken lessons from other countries' mistakes, and it set modest aims for its first interplanetary mission, which it billed as a technology demonstration. When Mangalyaan settled successfully into Mars orbit on 24 September, India joined the elite group of nations with the ambition and technical capability to explore the Solar System.

In his 43 years as an engineer and manager at the ISRO, Radhakrishnan has led a diverse set of projects, from developing remote-sensing satellites to setting up India's tsunami-warning system. The Mars mission was a gamble, but it caused less heartache than the ISRO's work on a cryogenic rocket engine that had failed during a launch in 2010 and finally succeeded this year. "The Mars mission was a slightly more joyous occasion," he says, while playing down his own role. "I was like a conductor of an orchestra."

The Mars mission has put the spotlight on Asia's space ambitions. India plans in the next three years to launch its second Moon mission, and China aims to bring lunar samples back to Earth by 2017.

India's success this year drew widespread applause. "This is good for India and its economy, demonstrating the ability to develop and implement high-technology enterprises," says Raymond Arvidson, a planetary scientist at Washington University in St. Louis, Missouri.

Radhakrishnan says that India's space plans should not be judged against those of other countries: "We are not racing with anyone. We are only racing with ourselves." But he will soon leave the race. Radhakrishnan will retire at the end of the year, leaving him free to pursue his love for classical South Indian singing and dancing. He has not had much time for that during the ISRO's hectic pursuit of Mars. ■



**PETE
FRATES**



STEM-CELL TESTER

An ophthalmologist injected hope into the stem-cell field during a troubled year.

BY DAVID CYRANOSKI

For an hour on Friday 12 September, Masayo Takahashi sat alone, calmly reflecting on the decade of research that had led up to this moment.

An ophthalmologist at the RIKEN Center for Developmental Biology (CDB) in Kobe, Japan, Takahashi was about to watch a sheet of epithelial cells that she had grown be transplanted into the back of a woman's damaged retina. She had made the cells from induced pluripotent stem (iPS) cells, which have been widely touted for their potential to generate genetically-matched tissue for treating a range of diseases. The transplant would be the first test of that promise in people, and therefore a major milestone for the stem-cell field. As she sat, Takahashi quietly considered all those who had helped her get to that point ("so many people — it would be like the credits rolling at the end of a movie"), and the scandal in the stem-cell field that had threatened to derail the project earlier in the year. "It was like a sacred hour," she says.

Takahashi had been trying to use stem cells to repair retinal damage for ten years — and trying to downplay hype about the cells for almost as long. Her work received a boost when, in 2006, stem-cell scientist Shinya Yamanaka at Kyoto University in Japan discovered how to make iPS cells, which are much easier to make than other human pluripotent cells. Collaborating with Yamanaka, Takahashi worked out how to turn the iPS cells into sheets of retinal epithelial cells. She then tested the resulting cells in mice and monkeys, passed regulatory hurdles, recruited patients, and practised growing cells from those patients. Finally, she

MAMI NAGAOKI/AP/PRESS ASSOCIATION IMAGES

STRUCTURE SOLVER

A biologist brought the cell's molecular machines into sharper focus.

BY EWEN CALLAWAY

Sjors Scheres is surrounded by ribosomes. A picture of one fills his computer screen, and thousands more are stuffed on his hard drive. His CV is studded with high-profile papers from this year showing some of the clearest images ever produced of these complex protein-making machines. So it is all the more surprising when Scheres, a structural biologist, says that he isn't all that interested in ribosomes. "It's all about the math," he says, with relish. "That's what my main contribution is."

That mathematics is helping to drive a revolution in structural biology. Once dominated by a method called X-ray crystallography, the field is now in the thrall of a technique called cryo-electron microscopy, or cryo-EM.

Scheres's calculations have led to software that transforms grainy cryo-EM images into exquisitely detailed pictures, allowing biologists to visualize molecular machines more easily and accurately than ever before.

Scheres started his PhD trying to get a portion of a gene-regulation protein to form tidy crystals — a prerequisite for X-ray crystallography, which involves pummeling the crystals with X-rays, then using the resulting diffraction patterns to deduce the protein's shape. But he abandoned the project when his protein, like so many others, defied crystallization. He was drawn instead to cryo-EM, in which a beam of electrons is used to visualize flash-frozen protein solutions. Three-dimensional structures are then created by merging electron micrographs taken from different angles. But at the time, the technique was known as 'blob-ology' because the images it produced were so patchy, Scheres says.

In 2010, when Scheres joined the Laboratory for Molecular Biology (LMB) in Cambridge, UK, microscopes were being developed that could detect electrons more efficiently and take snapshots of proteins at hundreds of frames per second. But Scheres knew that better computer programs would be needed to make sense of the flood of data, so he shut himself in his office to try to write one. "I didn't have a group. I was just programming," he says. The resulting software, named RELION, brought the blobs into focus: it did a much better job of marrying images

Ones to watch

2015

XIE ZHENHUA

CHINA'S TOP CLIMATE OFFICIAL

After this year's climate accord between the United States and China, Xie and the world's biggest greenhouse-gas polluter will be a focus of attention at climate talks.

ALAN STERN

PRINCIPAL INVESTIGATOR OF NASA'S NEW HORIZONS MISSION

Stern will be firmly in the spotlight in July when his mission becomes the first to reach Pluto. Just don't tell him it's not a planet.

JOANNE LIU

INTERNATIONAL PRESIDENT OF MÉDECINS SANS FRONTIÈRES (MSF)

MSF has shone in the global response to the Ebola epidemic, and Liu will be a big player in next year's efforts to end it.

BERNARD BIGOT

NOMINATED AS NEXT DIRECTOR-GENERAL OF ITER

Bigot wants to radically reform the troubled multi-billion-euro international project to build a huge reactor that would demonstrate the feasibility of fusion energy.

RICK HORWITZ

EXECUTIVE DIRECTOR, ALLEN INSTITUTE FOR CELL SCIENCE

As head of a new US\$100-million venture funded by philanthropist Paul Allen, Horwitz must push cell biology to a new frontier.

was ready to try the transplants in people with a common condition called age-related macular degeneration, in which wayward blood vessels destroy photoreceptors and vision. The transplants are meant to cover the retina, patch up the epithelial layer and support the remaining photoreceptors. Watching the procedure, "I could feel the tension of the surgeon," Takahashi says.

In the end, everything went smoothly — but Takahashi will not reveal whether it has been a success until a year after the transplant. She does say that the tissue seems to have maintained its brownish colour, a sign that it has not been attacked by the immune system. The patient, a woman in her 70s, had already lost most of her vision and is unlikely to get it back; but Takahashi's team is keen to see whether the transplant is safe and prevents further retinal deterioration.

Takahashi had planned to operate on six patients in an informal clinical study. But a law that went into effect in Japan last month opens the door to a fast-track formal trial that would move the technology, if successful, to open clinical use. She is now considering which path to take.

The transplant was a high point for the field after a major low. Earlier in the year, controversy over two stem-cell papers published in *Nature* and unrelated to Takahashi's research had enveloped the CDB. The papers, which reported a quick recipe for making pluripotent stem cells, were first lauded and then shunned after it emerged that some figures had

been manipulated. The spotlight fell on Haruko Obokata, the papers' first author, who continued to argue that the method worked. The episode took a tragic turn when Yoshiki Sai, who supervised Obokata at the CDB, killed himself in August. In the wake of the scandal, the centre was drastically restructured and its research budget was slashed.

As all this unfolded, Takahashi found her own work under intense scrutiny: she was accused of rushing the procedure in an effort to make money, and concerns were raised over whether the cells were safe. A month before the scheduled surgery, the health ministry suddenly announced that several new safety tests would be required. At times, Takahashi says, she felt "beaten".

Now upbeat, however, Takahashi is aiming to clear a much higher bar — transplanting layers of photoreceptors together with the epithelial sheets — to restore a small degree of vision to people with macular degeneration. The photoreceptors would have to make connections with neurons, something that Takahashi realizes will be a challenge. For that, she will use the ability to grow three-dimensional retinal tissue *in vitro* — a technique, she notes with sadness, that was pioneered by Sai.

Other scientists at the centre share the grief, and say that Takahashi's success was a welcome distraction. "It was definitely encouraging for all CDB people," says developmental biologist Masatoshi Takeichi, former director of the centre. ■

MRC LMB

into a three-dimensional molecular structure than did existing tools.

"We left him alone for a couple years, and he came up with all this beautiful software," says Venki Ramakrishnan, a molecular biologist at the LMB. Ramakrishnan had won the 2009 Nobel Prize in Chemistry for his work in determining the structure of the bacterial ribosome using X-ray crystallography. But it takes years to obtain such structures because ribosomes are made up of dozens of different proteins and RNA molecules. Cryo-EM offers a quicker route, and this year, Ramakrishnan collaborated with Scheres to produce detailed structures of yeast and human ribosomes. Now, his lab has converted almost exclusively to the new technology. "For us it's a perfect saviour," he says. "We can be defined by the biological questions, rather than what we can crystallize."

Scheres is now looking for more difficult structures to crack. He found one in a project with a team at Tsinghua University in Beijing, to determine the structure of γ -secretase, a protein implicated in Alzheimer's disease. The protein is relatively small and prone to movement, which blurs cryo-EM images — but Scheres has already produced one structure and is working on improvements. "It is kind of a boom time in cryo-EM," says Richard Henderson, a structural biologist at the LMB who helped to develop the new electron microscopes, "and Sjors deserves a lot of the credit for getting it going." ■



COMMENT

EPIDEMIOLOGY Could Ebola survivors help to shrink the epidemic? **p.323**



MARINE BIOLOGY Ed Ricketts, the taxonomist behind John Steinbeck's heroes **p.326**

CONSERVATION Protected areas empower communities and attract investment **p.329**

OBITUARY Martin L. Perl, Nobel-winning discoverer of tau lepton, remembered **p.330**



Defend the integrity of physics

Attempts to exempt speculative theories of the Universe from experimental verification undermine science, argue **George Ellis** and **Joe Silk**.

This year, debates in physics circles took a worrying turn. Faced with difficulties in applying fundamental theories to the observed Universe, some researchers called for a change in how theoretical physics is done. They began to argue — explicitly — that if a theory is sufficiently elegant and explanatory, it need not be tested experimentally, breaking with centuries of philosophical tradition of defining scientific knowledge as empirical. We disagree. As the philosopher of science Karl Popper argued: a theory must be falsifiable to be scientific.

Chief among the 'elegance will suffice' advocates are some string theorists. Because string theory is supposedly the 'only game in town' capable of unifying the four fundamental forces, they believe that it must contain a grain of truth even though it relies on extra dimensions that we can never observe. Some cosmologists, too, are seeking to abandon experimental verification of grand hypotheses that invoke imperceptible domains such as the kaleidoscopic multiverse (comprising myriad universes), the 'many worlds' version of quantum reality (in

which observations spawn parallel branches of reality) and pre-Big Bang concepts.

These unprovable hypotheses are quite different from those that relate directly to the real world and that are testable through observations — such as the standard model of particle physics and the existence of dark matter and dark energy. As we see it, theoretical physics risks becoming a no-man's-land between mathematics, physics and philosophy that does not truly meet the requirements of any.

The issue of testability has been lurking ►

ILLUSTRATION BY VASAVA

► for a decade. String theory and multiverse theory have been criticized in popular books^{1–3} and articles, including some by one of us (G.E.)⁴. In March, theorist Paul Steinhardt wrote⁵ in this journal that the theory of inflationary cosmology is no longer scientific because it is so flexible that it can accommodate any observational result. Theorist and philosopher Richard Dawid⁶ and cosmologist Sean Carroll⁷ have countered those criticisms with a philosophical case to weaken the testability requirement for fundamental physics.

We applaud the fact that Dawid, Carroll and other physicists have brought the problem out into the open. But the drastic step that they are advocating needs careful debate. This battle for the heart and soul of physics is opening up at a time when scientific results — in topics from climate change to the theory of evolution — are being questioned by some politicians and religious fundamentalists. Potential damage to public confidence in science and to the nature of fundamental physics needs to be contained by deeper dialogue between scientists and philosophers.

STRING THEORY

String theory is an elaborate proposal for how minuscule strings (one-dimensional space entities) and membranes (higher-dimensional extensions) existing in higher-dimensional spaces underlie all of physics. The higher dimensions are wound so tightly that they are too small to observe at energies accessible through collisions in any practicable future particle detector.

Some aspects of string theory can be tested experimentally in principle. For example, a hypothesized symmetry between fermions and bosons central to string theory — supersymmetry — predicts that each kind of particle has an as-yet-unseen partner. No such partners have yet been detected by the Large Hadron Collider at CERN, Europe's particle-physics laboratory near Geneva, Switzerland, limiting the range of energies at which supersymmetry might exist. If these partners continue to elude detection, then we may never know whether they exist. Proponents could always claim that the particles' masses are higher than the energies probed.

Dawid argues⁶ that the veracity of string theory can be established through philosophical and probabilistic arguments about the research process. Citing Bayesian analysis, a statistical method for inferring the likelihood that an explanation fits a set of facts, Dawid equates confirmation with the increase of the probability that a theory is true or viable. But that increase of probability can be purely theoretical. Because “no-one has found a good alternative” and “theories without alternatives tended to be viable in the past”, he reasons that string

theory should be taken to be valid.

In our opinion, this is moving the goalposts. Instead of belief in a scientific theory increasing when observational evidence arises to support it, he suggests that theoretical discoveries bolster belief. But conclusions arising logically from mathematics need not apply to the real world. Experiments have proved many beautiful and simple theories wrong, from the

“The consequences of overclaiming the significance of certain theories are profound.”

steady-state theory of cosmology to the SU(5) Grand Unified Theory of particle physics, which aimed to unify the electroweak force and the strong force. The idea that preconceived truths about the world can be inferred beyond established facts (inductivism) was overturned by Popper and other twentieth-century philosophers.

We cannot know that there are no alternative theories. We may not have found them yet. Or the premise might be wrong. There may be no need for an overarching theory of four fundamental forces and particles if gravity, an effect of space-time curvature, differs from the strong, weak and electromagnetic forces that govern particles. And with its many variants, string theory is not even well defined: in our view, it is a promissory note that there might be such a unified theory.

MANY MULTIVERSES

The multiverse is motivated by a puzzle: why fundamental constants of nature, such as the fine-structure constant that characterizes the strength of electromagnetic interactions between particles and the cosmological constant associated with the acceleration of the expansion of the Universe, have values that lie in the small range that allows life to exist. Multiverse theory claims that there are billions of unobservable sister universes out there in which all possible values of these constants can occur. So somewhere there will be a bio-friendly universe like ours, however improbable that is.

Some physicists consider that the multiverse has no challenger as an explanation of many otherwise bizarre coincidences. The low value of the cosmological constant — known to be 120 factors of 10 smaller than the value predicted by quantum field theory — is difficult to explain, for instance.

Earlier this year, championing the multiverse and the many-worlds hypothesis, Carroll dismissed Popper's falsifiability criterion as a “blunt instrument” (see go.nature.com/nuj39z). He offered two other requirements: a scientific theory should be “definite” and “empirical”. By definite, Carroll means that the theory says

“something clear and unambiguous about how reality functions”. By empirical, he agrees with the customary definition that a theory should be judged a success or failure by its ability to explain the data.

He argues that inaccessible domains can have a “dramatic effect” in our cosmic backyard, explaining why the cosmological constant is so small in the part we see. But in multiverse theory, that explanation could be given no matter what astronomers observe. All possible combinations of cosmological parameters would exist somewhere, and the theory has many variables that can be tweaked. Other theories, such as unimodular gravity, a modified version of Einstein's general theory of relativity, can also explain why the cosmological constant is not huge⁷.

Some people have devised forms of multiverse theory that are susceptible to tests: physicist Leonard Susskind's version can be falsified if negative spatial curvature of the Universe is ever demonstrated. But such a finding would prove nothing about the many other versions. Fundamentally, the multiverse explanation relies on string theory, which is as yet unverified, and on speculative mechanisms for realizing different physics in different sister universes. It is not, in our opinion, robust, let alone testable.

The many-worlds theory of quantum reality posed by physicist Hugh Everett is the ultimate quantum multiverse, where quantum probabilities affect the macroscopic. According to Everett, each of Schrödinger's famous cats, the dead and the live, poisoned or not in its closed box by random radioactive decays, is real in its own universe. Each time you make a choice, even one as mundane as whether to go left or right, an alternative universe pops out of the quantum vacuum to accommodate the other action.

Billions of universes — and of galaxies and copies of each of us — accumulate with no possibility of communication between them or of testing their reality. But if a duplicate self exists in every multiverse domain and there are infinitely many, which is the real ‘me’ that I experience now? Is any version of oneself preferred over any other? How could ‘I’ ever know what the ‘true’ nature of reality is if one self favours the multiverse and another does not?

In our view, cosmologists should heed mathematician David Hilbert's warning: although infinity is needed to complete mathematics, it occurs nowhere in the physical Universe.

PASS THE TEST

We agree with theoretical physicist Sabine Hossenfelder: post-empirical science is an oxymoron (see go.nature.com/p3upwp and go.nature.com/68rijj). Theories such as quantum mechanics and relativity turned

out well because they made predictions that survived testing. Yet numerous historical examples point to how, in the absence of adequate data, elegant and compelling ideas led researchers in the wrong direction, from Ptolemy's geocentric theories of the cosmos to Lord Kelvin's 'vortex theory' of the atom and Fred Hoyle's perpetual steady-state Universe.

The consequences of overclaiming the significance of certain theories are profound — the scientific method is at stake (see go.nature.com/hh7mm6). To state that a theory is so good that its existence supplants the need for data and testing in our opinion risks misleading students and the public as to how science should be done and could open the door for pseudoscientists to claim that their ideas meet similar requirements.

What to do about it? Physicists, philosophers and other scientists should hammer out a new narrative for the scientific method that can deal with the scope of modern physics. In our view, the issue boils down to clarifying one question: what potential observational or experimental evidence is there that would persuade you that the theory is wrong and lead you to abandoning it? If there is none, it is not a scientific theory.

Such a case must be made in formal philosophical terms. A conference should be convened next year to take the first steps. People from both sides of the testability debate must be involved.

In the meantime, journal editors and publishers could assign speculative work to other research categories — such as mathematical rather than physical cosmology — according to its potential testability. And the domination of some physics departments and institutes by such activities could be rethought^{1,2}.

The imprimatur of science should be awarded only to a theory that is testable. Only then can we defend science from attack. ■

George Ellis is professor emeritus of applied mathematics at the University of Cape Town, South Africa. **Joe Silk** is professor of physics at the Paris Institute of Astrophysics, France, and at Johns Hopkins University in Baltimore, Maryland, USA.
e-mails: george.ellis@uct.ac.za; silk@iap.fr

1. Woit, P. *Not Even Wrong* (Cape, 2006).
2. Smolin, L. *The Trouble with Physics* (Penguin, 2006).
3. Baggott, J. *Farewell to Reality* (Constable, 2013).
4. Ellis, G. F. R. *Sci. Am.* **305**, 38–43 (2011); available at <http://go.nature.com/27p6oe>.
5. Steinhardt, P. *Nature* **510**, 9 (2014).
6. Dawid, R. *Phil. Sci.* **73**, 298–332 (2007).
7. Ellis, G. F. R. *Gen. Rel. Grav.* **46**, 1619 (2014).



JOHN MOORE/GETTY

Ebola survivors Zaiyay Mulbah (left), a former money changer, and Mark Jerry, previously a delivery driver, are working as nurses' assistants at a Liberian Ebola centre.

Mobilizing Ebola survivors to curb the epidemic

Scaling up the recruitment of individuals who have recovered from infection deserves urgent consideration, argue **Joshua M. Epstein**, **Lauren M. Sauer** and colleagues.

Multiple governments and non-governmental organizations have called on health-care personnel the world over to help control West Africa's Ebola outbreak; these include Médecins Sans Frontières (MSF), the World Health Organization (WHO) and United Nations children's charity UNICEF. But the

demand for labour far exceeds the supply¹. UN estimates, which may be low, suggest that approximately 5,000 international medical, training and support personnel are needed in the coming months.

While foreign assistance must continue, a nascent local strategy is a candidate for broad adoption. We call it MORE, for ►



Ebola survivors are assisting in World Health Organization response efforts, which could be expanded.

► Mobilization of REcovered individuals. The idea is simple: those who have recovered from Ebola could be engaged to reduce transmission, helping to bring the epidemic under control.

Examples of the approach can be seen in Sierra Leone, Guinea and Liberia. For instance, the UN is training survivors to support children who have had contact with infected individuals and are within Ebola's 21-day incubation window (the time it takes to develop symptoms after being infected with the virus). MSF is similarly employing survivors to work in their Ebola treatment units in Guinea and Liberia.

There are uncertainties about the ultimate size of this cadre and, crucially, about the immunity of recovered responders to reinfection, both immediately and in the longer term (because immunity may wane). Nonetheless, the potential of MORE to shift the epidemic's dynamics makes its consideration imperative.

RECOVERED RESPONDERS

So far, Ebola has infected an estimated 16,000 individuals in Liberia, Sierra Leone and Guinea. Current estimates suggest that in West Africa, roughly 50% of people who contract Ebola will die². This would leave a substantial pool of survivors, totalling perhaps 8,000 people by the end of the year. In the longer term, this could prove to be a much larger number. Indeed, the larger the epidemic, the bigger this pool becomes.

The worst-case projections of the US Centers for Disease Control and Prevention, for Sierra Leone and Liberia only, range from 500,000 to 1.4 million cases

of Ebola by January 2015 (ref.3). Owing to various methodological limitations (set forth earlier by one of us, J.M.E., see go.nature.com/86kpyw), these projections are proving to be much too high. But even if the lower of these estimates turns out to be an order of magnitude too high, there could ultimately be 50,000 cases. If 50% survive, this is a pool of 25,000. If we assume that 75% of survivors would be too young, too old, too ill or too traumatized to be recruited, the available cadre could still number in the thousands (see go.nature.com/kbx4el).

There are limited confirmatory data on protective immunity to Ebola in humans. But researchers generally agree that the evidence is pointing towards survivors being immune to reinfection. Thus far, there has not been a single reported case of a person who recovered from *Zaire ebolavirus* (the lineage of the current outbreak) becoming reinfected. This, and evidence from animal studies, suggests that people may have protective immunity following recovery.

RISK LEVEL

At worst, recovered responders would have the same level of risk as the general population, in which case they would need to use the same personal protective equipment (PPE) as other responders. At best, they would have high protection through conferred immunity.

In the latter event, recovered responders could operate with much less onerous PPE than current health-care workers. They would require only the training and protective equipment (medical gloves, face shield and goggles) used to minimize the transmission of more familiar blood-borne pathogens such as HIV. This would allow them to have much more extensive contact with patients than Ebola PPE normally affords. Generally, providers in full Ebola PPE work only two-hour shifts to avoid overheating (see go.nature.com/hsk4v5).

Recovered individuals can be trained to perform many important response functions (see 'Responder roles'). Some of these are beyond palliative, and may have a direct impact on disease transmission, changing the course of the epidemic itself. Such activities include isolating suspected patients from uninfected community members

EBOLA CONTROL

Reversing the epidemic

In classical epidemiology, susceptible people (*S*) bump into infected ones (*I*) as in a perfectly mixed bowl. That is, if β is the transmission probability per contact between these pools, the epidemic grows at rate βSI .

To model the impact of the Mobilization of REcovered individuals (MORE) strategy, we let Z_t denote the recovered proportion of the population t days into the epidemic, and k denote the fraction of recovered people who are deployed to reduce transmission. This multiplies the classical growth rate above by $(1 - kZ_t)$, which one might interpret as reducing β .

The reproductive number, R_t , is the average number of primary infections produced by a single infected individual

dropped into the population on day t . If $R_t > 1$ the epidemic is growing, whereas if $R_t < 1$, it is declining. So, $R_t = 1$ is the epidemic threshold. Letting R_t^M and R_t^C denote, respectively, the reproductive numbers in the MORE and classical models (including deaths), it follows that:

$$R_t^M = R_t^C (1 - kZ_t).$$

So, if k (the mobilized fraction of the recovered) exceeds zero, MORE reduces the reproductive number. And crucially, if $R_t = 1$ or is hovering above it, mobilized survivors could tip the epidemic into fading out. More-realistic models, with social networks rather than perfect mixing, could reveal stronger effects.

RESPONDER ROLES

Many tasks that could help to bring the Ebola epidemic under control require only limited training. Recovered individuals could be screened and assigned functions on the basis of their preferences and capabilities.

IMMUNITY NOT ESSENTIAL

Low or moderate skill

Emergency response management operations
Family support
Stigma education
Contact tracing
Nutrition services

High skill

Hygiene education
Burial education

IMMUNITY ESSENTIAL

Low or moderate skill

Patient screening
Waste management
Burial monitoring
Physical labour (moving supplies, building treatment tents, for example)
Housekeeping
Patient morale

High skill

Patient transport
Early identification of infected

and safely transporting them to treatment centres early in the course of infection, when viral loads and contagiousness are lowest. Performing simple tasks that do not require extensive training — such as giving people food and water, helping them to shower or feeding infants — would free specialized health-care workers to concentrate on more sophisticated clinical tasks. Recovered responders could also perform duties such as waste disposal and decontamination in high-risk areas ('hot zones') and ensure safe burial practices, all of which reduce the spread of the virus.

As well as reducing Ebola transmission, MORE could enhance West African health infrastructure. It will generate a form of human capital that will continue to facilitate routine health care and the early warning of Ebola recurrence when the international presence declines. Furthermore, recovered responders would be important allies in any vaccination campaign.

TIPPING POINT

Most importantly, mobilizing the recovered to reduce transmission could 'tip' the epidemic into decline. Specifically, a central idea in epidemic modelling is the reproductive number of the disease, denoted by R_t (see 'Reversing the epidemic'). This is interpreted as the average number of primary infections produced by a single infected individual dropped into the population at a particular time, t . If everyone is already ill there is no one to infect, R_t is effectively zero.

By contrast, the very first infectious person introduced into a dense, uninfected

population might transmit the disease to many, so the reproductive number at this time ($t=0$) would be high. If R_t is greater than 1, the epidemic is growing: each infected person is converting more than one susceptible person into another infective person. If R_t is less than 1, the epidemic is shrinking. Therefore, the state at which R_t is equal to 1 can be considered a tipping point. Above this point, the disease takes off. Below it, the epidemic dies away. MORE could reduce R_t ; in countries currently close to the tipping point, such as Liberia, the strategy could bring R_t below 1 and keep it there.

The World Bank is poised to spend almost US\$500 million on foreign response for West Africa. With the economies of the most affected countries severely strained, this external support is crucial. By comparison, a small investment could establish a complementary standing cadre of local recovered responders. (Liberia's nurses, for example, are paid roughly \$10 per day.) This strategy would create local jobs paying a fair wage in places where these are in short supply.

TACKLING STIGMA

A serious concern is that Ebola survivors are stigmatized. This was true of HIV in the early stages of the AIDS epidemic. There, stigma and social marginalization were successfully reduced with intensive educational campaigns and the support of national and international leaders⁴. Similar tactics can be employed here, but much earlier, and with potentially dramatic effects. In fact, local survivor support groups and other efforts are already leading the way (see go.nature.com/t1xv4f).

MSF has begun providing 'certificates of recovery' to survivors, as a means of allaying fears. Community volunteers in Sierra Leone and Liberia are working to combat misplaced fear: by visiting the neighbours of survivors before their return home; by broadcasting on radio and television; and by physically embracing survivors themselves. Social-media campaigns are reinforcing the message that people are not defined by their disease — among these are the 'I am a Liberian, not a virus!' and the 'I survived Ebola' campaigns.

If these campaigns are effective, people infected with — or at risk of — Ebola may be particularly responsive to local survivors, who share their cultures, customs and language. In addition, recovered responders may themselves benefit from the work, improving their own psychological recovery.

High levels of illiteracy could be a concern insofar as they preclude highly specialized training. But for many of the tasks that we have highlighted, only limited training is required. People could be screened⁵ and assigned functions on the basis of their

preferences, aspirations, capabilities and work experience.

Crucial to the success of the MORE strategy is implementation on a broad scale. Although several groups are sporadically using survivors, the designation of a central body to implement and manage MORE would facilitate an efficient broadening of the approach. The WHO is currently the lead organization for the international Ebola virus disease response, and it has direct relationships with local governments, ministries of health and response organizations working on the ground. An alternative lead organization, MSF, is already working in this arena. MSF has substantial experience in managing volunteers in clinical environments and has

"An urgent priority is to establish precisely which individuals indeed have protective immunity."

comprehensive training programmes for various skill levels. A joint initiative of the WHO and MSF might be the best option for coordinated and effective implementation of MORE.

An urgent priority is to establish precisely which individuals indeed have protective immunity. This requires the identification of markers associated with immunity in the blood and serum of survivors⁶. Large-scale *in vivo* epidemiological studies will be important in solving this problem. Only then will we know the full potential of MORE. Meanwhile, recovered responders could perform tasks for which immunity is not essential, or use PPE where it is, as immediate steps in this promising direction. ■

Joshua M. Epstein is professor of emergency medicine and director of the Center for Advanced Modeling at Johns Hopkins University, Baltimore, Maryland, USA.

Lauren M. Sauer is a research associate in the Department of Emergency Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, USA. **Julia Chelen, Erez Hatna, Jon Parker, Richard E. Rothman, Lewis Rubinson.**
e-mail: jepste15@jhmi.edu

1. WHO Ebola Response Team *N. Engl. J. Med.* **371**, 1481–1495 (2014).
2. World Health Organization Ebola Response Roadmap Situation Report: 26 November 2014 (WHO, 2014).
3. Meltzer, M. I. et al. *Morbidity and Mortality Weekly Report* **63**, 1–14 (2014).
4. Mahajan, A. P. et al. *AIDS* **22**, (Suppl. 2) S67–S79 (2008).
5. USAID Community Health Worker Assessment and Improvement Matrix (USAID, 2011); available at <http://go.nature.com/cm1csu>
6. Bellan, S. E., Pulliam, J. R. C., Dushoff, J. & Meyers, L. A. *Lancet* **384**, 1499–1500 (2014).

A full list of author affiliations accompanies this article online at go.nature.com/eqkw4j.

IN RETROSPECT

Between Pacific Tides

Aaron Hirsh celebrates the 75th anniversary of the marine-biology classic by Ed Ricketts, the bohemian scientist who inspired John Steinbeck.



Marine biologist Ed Ricketts holding a Humboldt squid outside his laboratory in Monterey, California.

The heroes in a number of John Steinbeck's novels look alike: they are virile yet gentle men, full-blooded but also philosophical. There is a reason for the family resemblance among these characters, from Jim Casy in *The Grapes of Wrath* to Doctor Winter in *The Moon is Down*: they are all based on a person whom Steinbeck loved and admired. His name was Ed Ricketts, and he wrote books of his own. Not novels, but works of science and philosophy.

One of them, *Between Pacific Tides*, was a guide to the marine invertebrates of North America's Pacific shore, illustrated with black-and-white photographs by Jack Calvin. And although that might sound like a rather sober and orthodox piece of work, the vital and avid personality that captivated Steinbeck also made this particular field guide unique, compelling and ultimately enduring. Even today, 75 years after the book's publication, every marine biologist knows just where to reach for his or her own dog-eared and water-warped copy.

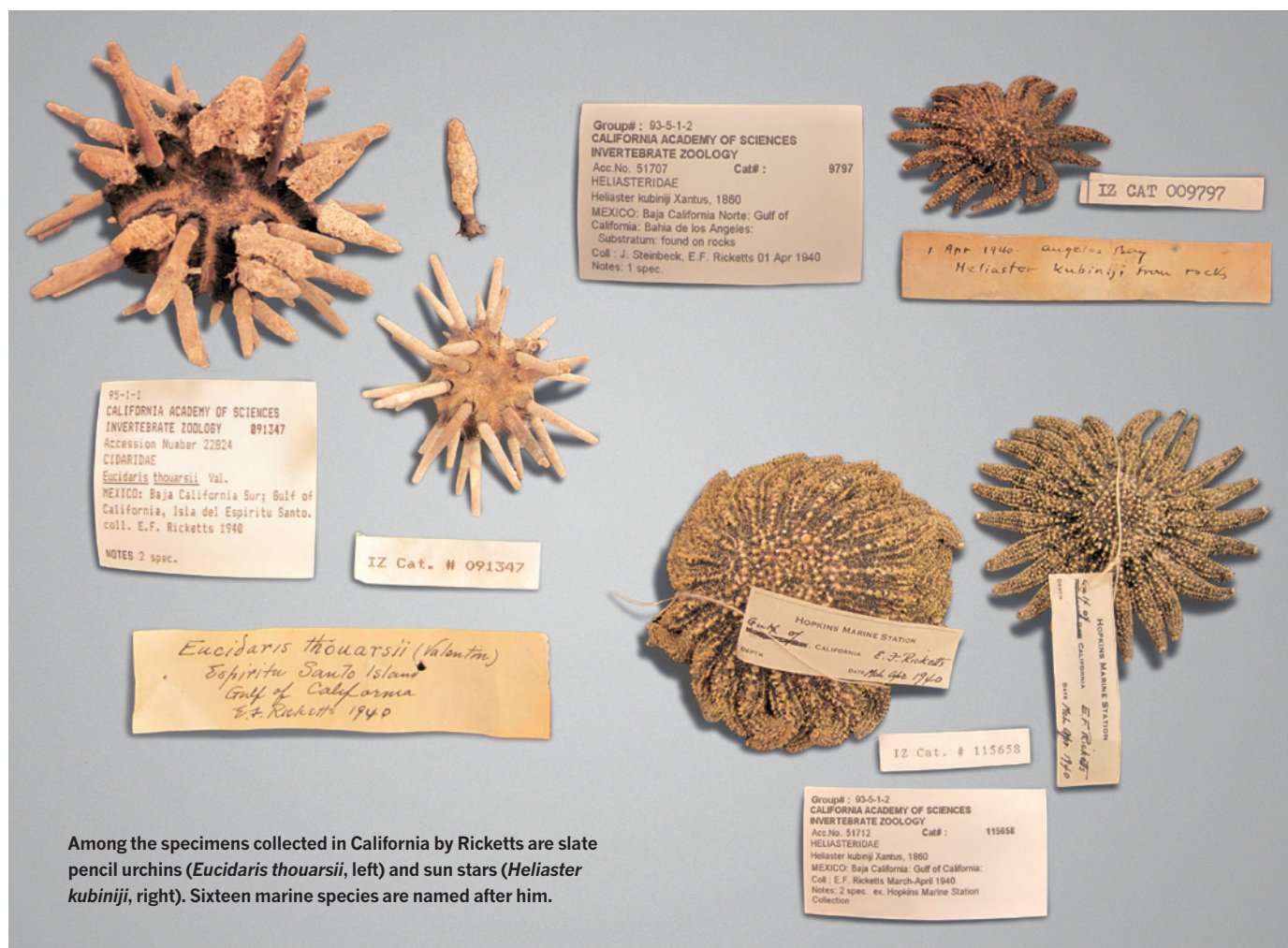
To get a sense of Ricketts, we can look to the Steinbeck hero who comes closest to being straight biography. Doc, from the 1945 novella *Cannery Row*, dresses like a vagrant but talks like a prophet. In mind and body alike, zealous appetites are balanced by overflowing generosity. He recites with equal verve the spiritual verse of eighth-century Chinese poet Li Po and the mysterious properties of marine invertebrates: the writhing brittle stars and the ravishing nudibranchs, the flatworms and the ribbon worms, the impervious limpets and the tide-pool shrimp so transparent you can see their tiny hearts. Doc is a merchant in these exotic beings. He hires tramps to collect them, then mails the creatures off to classrooms or laboratories.

All this is an accurate depiction of Ricketts and his world. Real, too, was Doc's zoological supply company on Cannery Row in Monterey, California: a weathered wooden structure where intellectuals, prostitutes and drunks convened to discuss philosophy and art amid books, pickled animals and a

tank of live dog-fish. Ricketts was a Socrates to these bibulous symposia, making a profound

Between Pacific Tides
EDWARD F. RICKETTS AND
JACK CALVIN
Stanford University Press:
1939.

RALPH BUCHSBAUM/VICKI BUCHSBAUM PEARSE AND CALIFORNIA HISTORY ROOM, MONTEREY PUBLIC LIB.



Among the specimens collected in California by Ricketts are slate pencil urchins (*Eucidaris thouarsii*, left) and sun stars (*Heliaster kubiniji*, right). Sixteen marine species are named after him.

impression not only on Steinbeck, but also on fellow novelist Henry Miller and the young Joseph Campbell, just finding his way into mythography.

One scene in Cannery Row points precisely to what made *Between Pacific Tides* different from the other books of its day. Doc sits on the shore with Hazel, the drifter he has hired to collect starfish. Looking at stink bugs crawling on the ground, Hazel asks, "What they got their asses up in the air for?" Doc replies: "They're very common animals and one of the commonest things they do is put their tails up in the air. And in all the books there isn't one mention of the fact that they put their tails up in the air or why."

A number of writers have observed that what made *Between Pacific Tides* revolutionary was that its organization is ecological rather than taxonomic: it categorizes animals according to habitat, not phylum or family. But the organization is also what you might call subjective or experiential: the order of presentation, and the information the text offers, anticipates exactly what a novice — someone like Hazel, just arriving at the shore — would notice and wonder about.

The book begins at the uppermost zone

where a flood tide's waves barely splash our shoes. And the first creatures we encounter there are the ones we really would notice: the familiar things, such as periwinkles, and the teeming ones, like rock lice. Only then are our eyes directed to the rarer animals, culminating in a special reward for our sustained attention: the giant owl limpet, *Lottia gigantea*.



An expedition by John Steinbeck and Ricketts formed the core of their book *Sea of Cortez*.

So it goes in each zone, as Ricketts leads us deeper into the intertidal: common to rare; familiar to exotic; obvious to hidden. And about each animal we are told not just a Latin name, but something to grab hold of, something intriguing — that *L. gigantea*, for instance, changes sex from male to female as it grows, and that it defends a territory, pursuing and bulldozing off any invaders. (Who knew that a limpet could pursue?) The book's perspective is subjective and experiential in another way, as well: Ricketts is unabashed about sharing his sympathetic inference of animal experience. About hermit crabs, for instance, he says: "when they are not busy scavenging or love-making, the gregarious 'hermits' fight with tireless enthusiasm tempered with caution."

Love-making? Enthusiasm? These are not exactly scientific terms, so perhaps it is not surprising that the book perplexed certain academic scientists. Reviewing the pre-publication manuscript, the director of Stanford University's Hopkins Marine Station in Pacific Grove, California, deplored the book's organization and reminded the publisher that Ricketts, who had never taken a university degree, was not, after all,

“a professional zoologist”.

But what made the book “unscientific” — the prioritization of subjective experience — is exactly what made it engaging and enduring. It is hard to maintain that the book was revolutionary in a scientific sense, because a lineage of ecological thinkers, from Alexander Von Humboldt to Charles Darwin to George Bird Grinnell, had already recognized that interactions between species are key determinants of abundance and geographic range. What was truly new — and a bit magical — was that Ricketts took a mind-numbing quantity of cold, hard facts, stitched them together, added the spark of his own avid experience, and something stirred to life: a tide pool, in the middle of which stands the reader, newly awake.

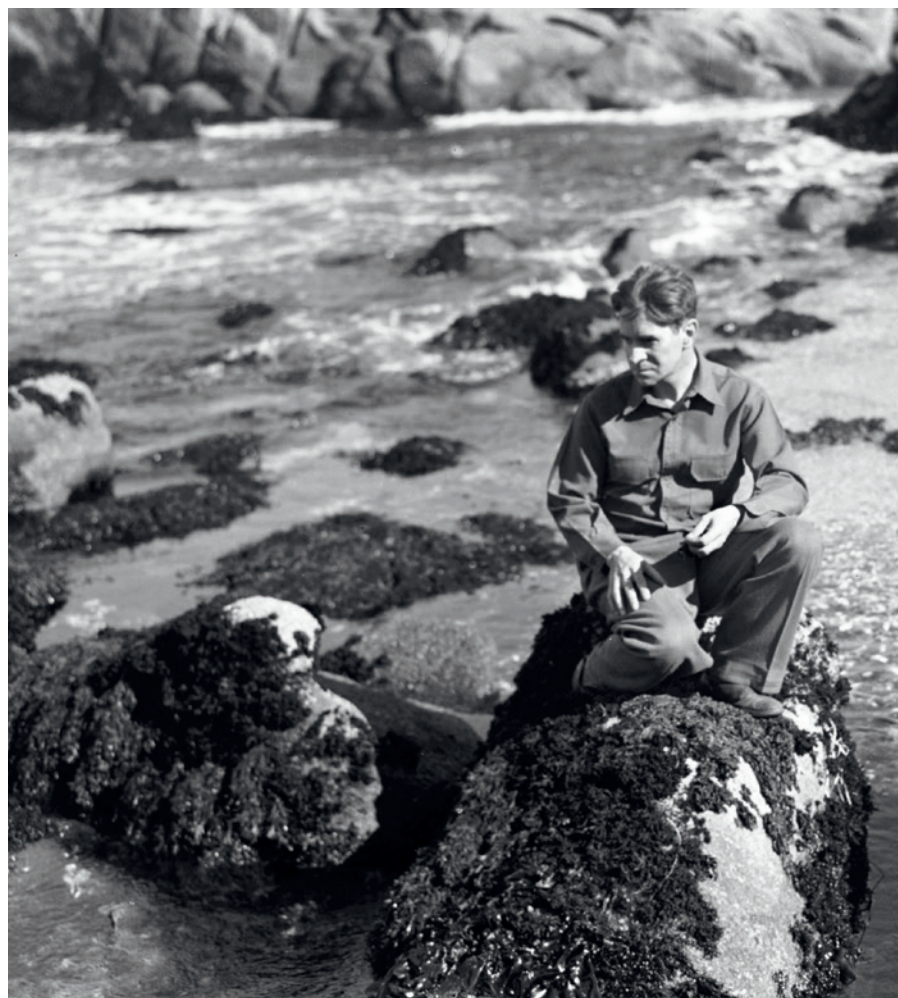
COMMUNING WITH NATURE

In this respect, the book is a unique creation of Ricketts’ personality and passion. As interested as he was in nature, Ricketts was more concerned with what it was like to be a human being in nature. While he was working on *Between Pacific Tides*, he also wrote three essays about what he called “breaking through”. They are literary and philosophical works, and they articulate his ambition not just to study the world objectively, from a distance, but to dive in — to connect, to sympathize. Even, you could say, to commune.

This impulse shows through the pages of *Between Pacific Tides* not just in amusing inferences about animal passions, but in other ways, too: Ricketts keeps smelling things (nudibranchs smell fruity, the fish called blennies too much like defunct kelp) and tasting things (owl limpets are delicious, nudibranchs, despite the smell, noxious), and generally digging his way into life and living. Steinbeck reported that he once saw Ricketts crawl inside a rotting basking shark to retrieve the liver for study.

Ricketts’ most ambitious and explicit effort to unify his scientific investigations with his more philosophical and spiritual quest was a book he wrote in collaboration with Steinbeck. *Sea of Cortez* is a weird and intriguing amalgam of literature, philosophy and science. The first half is a narrative about Steinbeck and Ricketts’ expedition together into the Gulf of California, the second half an illustrated catalogue of the animals they collected, including notes on distribution and abundance. Decades later, the book would prove key in documenting long-term ecological changes in the gulf, but at the time of its publication, in 1941, it was a flop.

“Steinbeck reported that he once saw Ricketts crawl inside a rotting basking shark to retrieve the liver for study.”



Ricketts among the tide pools on California’s Pacific shore in the late 1940s.

In 1948, Ricketts was killed when his car was hit by a train. A year later, the Viking Press secured Steinbeck’s permission to excise the zoological catalogue and reissue the narrative half of the book as *The Log from the Sea of Cortez*, with Steinbeck as sole author. Ricketts’ personal journal from the expedition, as well as his three philosophical essays, were first published in 2006 as part of the collection *Breaking Through* (University of California Press). As the editor of the collection points out, the texts reveal that whole paragraphs of the narrative now published under Steinbeck’s name were in fact only edited by him. They were written by Ricketts.

One of the lovelier bits of natural history that we learn in *Between Pacific Tides* pertains to the orange-and-white nudibranch *Triopha catalinae*. The creature, Ricketts wrote, “can be seen crawling upside down suspended from the underside of the air-water surface film of pools”. The image reminds me of Ricketts himself: he moved on the surface tension between two very different realms, science on one side, literature and philosophy on the other. It is such a delicate place to live, so microscopically thin;

even small disturbances at the surface send *Triopha catalinae* sinking to the bottom.

But Ricketts’ legacy brings to mind another trick of certain invertebrates: when you cut them into pieces, each bit grows up into a being of its own. In some ways, Ricketts never received the recognition he deserved — his essays unpublished during his lifetime, his most ambitious book cut apart. Yet his ideas and identity nevertheless proliferated, in various guises, in marine ecology and in mid-century intellectual culture. When Joseph Campbell writes of the mythic hero’s connection with animal powers, or when the hero of *The Grapes of Wrath* delivers a speech on social organization, or even when Henry Miller writes of sex as transcendence, the diverse descendant lineages of Edward F. Ricketts are propagated. As Steinbeck wrote: “He taught everyone without seeming to.” ■

Aaron Hirsh is a writer and biologist at the University of Colorado Boulder and the Vermilion Sea Institute. His most recent book is *Telling Our Way to the Sea: A Voyage of Discovery in the Sea of Cortez*. e-mail: aaron.hirsh@colorado.edu

Correspondence

Prescient words on comets and life

The landing of the Philae probe on comet 67P/Churyumov–Gerasimenko last month has led to speculation that comets might have delivered the building-block elements of life to Earth — an idea anticipated by the French astronomer Camille Flammarion more than a century ago in his 1880 book *Astronomie populaire*.

Flammarion wrote of comets: “Their importance would be much greater still if they should be found to carry in them the first combinations of carbon, for it is probable that it was by these combinations that vegetable and animal life commenced on the earth and the other planets and thus these vagrant bodies might be the sowers of life on all the worlds!”

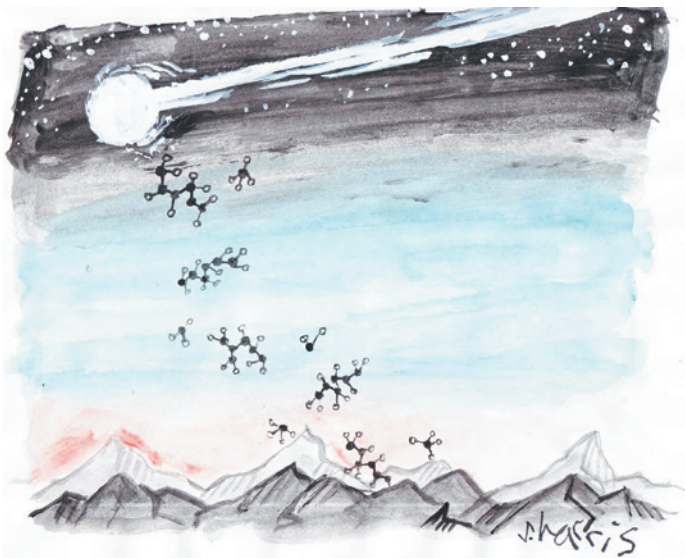
Milton Wainwright *University of Sheffield, UK*
m.wainwright@sheffield.ac.uk

Pool resources for protected areas

Protected conservation areas face huge challenges globally (see J. E. M. Watson *et al.* *Nature* **515**, 67–73; 2014). But examples that are effectively funded and managed can be found in Namibia and in the Brazilian Amazon. In our view, these models are so successful that they could be adapted and replicated around the world.

In Namibia, the Ministry of Environment and Tourism awards exclusive tourism concessions to communities that are next to protected areas and have formed conservancies. This attracts millions of dollars in infrastructure investment, empowering communities with economic activity and employment opportunities and creating strong incentives to live with and protect wildlife.

The Amazon Region Protected Areas programme safeguards a staggering 15% of the Brazilian



Amazon. Funded through an innovative partnership of public and private donors, it has secured US\$215 million to cover costs over the next 25 years. Brazil is gradually stepping up its own contributions to ensure full and permanent funding.

Jon Hoekstra, Meg Symington *World Wildlife Fund, Washington DC, USA*

Chris Weaver *WWF-Namibia, Windhoek, Namibia*
jon.hoekstra@wwfus.org

Research agency will lose autonomy

On 1 January 2015, a large new government office will take over Hungary's research-grant agency for basic science, OTKA. This will assume all budget management for research, development and innovation — destroying what the European Science Foundation has described as the agency's “high degree of political autonomy”.

A report by the foundation in November hailed OTKA as “the crown jewel of Hungary's R&D system; it is a professionally managed research council, whose procedures conform to the highest international standards. For several years, it has been in a constant process of improving its approaches and instruments, and it is obvious that OTKA

will continue to do so”. This view accords with that of most Hungarian scientists.

You note that “in Hungary, where the pluralism is under threat, the writing is on the wall” (see *Nature* **515**, 7–8; 2014). No scientist in Hungary expected that this prophecy would come true so soon.

András Váradi *Institute of Enzymology, RCNS, Hungarian Academy of Sciences, Hungary*

János Kertész *Central European University, Hungary*
varadi.andras@ttk.mta.hu

Flood resilience a must for delta cities

Conventional methods of flood protection such as levees are no longer adequate against the increased risk of flooding in Asian delta cities. We call for a multipronged approach that focuses on long-term, sustainable solutions to increase these cities' resilience to flooding (see also L. Giosan *et al.* *Nature* **516**, 31–33; 2014).

In October, Ho Chi Minh City in Vietnam experienced record-breaking flood levels in the Saigon River for the fifth consecutive year and for the eighth time in the past decade. Among the contributing factors are massive urban development, reduced

river-storage capacity, land subsidence from unregulated groundwater extraction, extreme storm events and rising sea levels. Similar disruptions in Beijing, Jakarta and Manila have also led to catastrophic floods.

Several strategies exist to increase resilience against flooding. These include developing urban infrastructure to decrease the effects of extreme rainfall (for example, by incorporating sustainable living green roofs and making pavements permeable); building in harmony with natural-systems dynamics, as in the Room for the River (go.nature.com/hqjld5) and Sand Motor (go.nature.com/e24ecq) projects in the Netherlands; and incorporating flood-risk forecasts for downstream urban areas into reservoir management.

Ruben Dahm* *Deltares, Delft, the Netherlands*
ruben.dahm@deltares.nl

*On behalf of 4 correspondents (see go.nature.com/qmy9vg for full list).

What football can teach science

One solution to the challenges posed by voluntary peer review (M. Arns *Nature* **515**, 467 (2014) and see *Nature* **515**, 480–482; 2014) might be to create a professional, independent body of reviewers that could be for hire by journals — rather like the professional referees used in football.

These reviewers could be funded by contributions from research councils, charities and end-users — namely, scientific journals and funding bodies.

Such a system could put an end to rigging scandals and to poor-quality, unprofessional or biased peer review, as well as improving the speed and consistency of the refereeing process. It could even offer stable employment for thousands of PhD graduates.

Arturo Sala *Brunel University London, Uxbridge, UK*
arturo.sala@brunel.ac.uk

Martin L. Perl

(1927–2014)

Discoverer of the tau lepton subatomic particle.

The tau lepton, a subatomic particle uncovered by Martin Lewis Perl, was one of the biggest surprises in elementary particle physics in recent decades. Perl discovered this third type of lepton (the other two types are the lighter electrons and muons) at a time when there was no experimental evidence for its existence or any theoretical indication that a third ‘family’ of particles should exist.

Perl, who died on 30 September at the age of 87, was born in 1927 in Brooklyn, New York, to Jewish immigrants from Poland. Through determination and hard work, his father had established a printing and advertising company that sustained the family throughout the depression of the 1930s. Perl once reflected that his parents’ high expectations — they demanded that he achieve A grades in every course — “was good training for research, because large parts of experimental work are sometimes boring or involve the use of skills in which one is not particularly gifted”.

Even though he graduated from high school at just age 16 and received a medal for his achievements in physics, Perl never thought of becoming a scientist. Neither he nor his family thought it was possible to make a living as a physicist. Perl decided instead to become a chemical engineer. His studies at the Polytechnic Institute of Brooklyn were interrupted by military service during the Second World War, but he completed his bachelor’s degree in 1948.

After graduating, Perl joined General Electric, where he worked in the electron-tube division. To develop and improve the company’s production process for vacuum tubes, at the time used in appliances such as televisions and radios, Perl needed to understand how the electron vacuum tube worked. He started taking physics courses at the Union College in Schenectady, New York, and it was here that he realized where his real interest lay.

In 1950, Perl left industry to start a PhD at Columbia University in New York, under the supervision of physicist Nobel laureate Isidor Isaac Rabi. The lessons that Perl took from Rabi — the importance of working on fundamental problems, choosing your own research problems, getting the right answers and checking them thoroughly before publishing — guided him throughout the rest of his career.

After a research and teaching job at the



LINDA A. CICERO/STANFORD NEWS SERVICE

University of Michigan in Ann Arbor, Perl got his first opportunity to think seriously about high-energy experiments on charged leptons when he was offered a position at the yet-to-be-built Stanford Linear Accelerator Center (SLAC) in Menlo Park, California. He moved to SLAC in 1963, and in December 1975, he and his colleagues published a paper entitled ‘Evidence for Anomalous Lepton Production in e^+e^- Annihilation’ (M. L. Perl *et al. Phys. Rev. Lett.* **35**, 1489; 1975). It was not until the end of 1979, however, that the discovery of the tau lepton was finally verified.

Until this point, the prevailing view among physicists had been that only two types of lepton existed: electrons and muons. The tau lepton is more than 3,000 times heavier than an electron and is highly unstable. Its discovery transformed the expectations of fundamental particle physics and paved the way for the discoveries of other elementary particles, including the tau neutrino and the bottom and top quarks. In 1995, Perl shared the Nobel Prize in Physics with Frederick Reines, who received his share for his part in the detection of the neutrino, another component of matter.

Martin’s enthusiasm for fundamental physics was contagious. Once, while working as a PhD student in his laboratory, I found an anomaly in our data that suggested the existence of a new fractionally charged particle. Martin, who treated any research he was involved in very seriously, immediately cancelled his trip to a scheduled conference

and stayed with our group for a few days. We pored over the data until we worked out — with some disappointment — that the anomaly was most likely an artefact of our experimental set-up.

As well as fundamental physics, Martin loved building mechanical devices and electrical instruments. When I and some other students were constructing an apparatus for an experiment, he was so curious and enthusiastic that he would frequently stop by to watch and learn about our progress. Martin always rewarded independent thinking. But while he helped his students to follow their own ideas, he also taught them to be realistic about what was possible — and to move on from a problem if they failed to make progress. His teachings served us both in the lab and later in life.

Martin had exceptionally high standards. He was a creative researcher who never chased honours, titles or respect, although respect always chased him. Everyone who knew him was impressed by his simplicity and honesty, summed up in the words he wrote at the end of his Nobel biography: “It was good fortune ...” ■

Valerie Halyo is a visiting scholar in experimental high-energy physics at Princeton University, New Jersey, USA. She earned her PhD under the supervision of Martin L. Perl at Stanford University in California from 1997 to 2001. e-mail: valerieh@princeton.edu

Better chemistry through radicals

An iron catalyst has been developed that mediates bond formation between a wide range of alkene reactants, opening up short synthetic routes to compounds that were previously accessible only through arduous pathways. [SEE ARTICLE P.343](#)

STEVEN L. CASTLE

Reactions that form carbon–carbon (C–C) bonds are essential for synthesizing complex organic molecules from simple, inexpensive precursors. The value that such molecules have as pharmaceuticals, agrochemicals and materials makes these reactions essential to the practice of organic synthesis. Most classical methods for generating C–C bonds rely on reagents that are either strongly basic or strongly acidic, and some require high temperatures to proceed. Such ‘harsh’ reaction conditions are incompatible with many functional groups — the groups of atoms responsible for the properties and reactivity of molecules. Functional-group incompatibilities are a major nuisance, because they force chemists to design synthetic routes to target molecules that are circuitous rather than direct. In an exciting development reported on page 343 of this issue, Lo *et al.*¹ have developed a C–C bond-forming reaction that provides a promising solution to this problem.

Since the 1970s, several C–C bond-forming reactions catalysed by transition-metal complexes have been developed that use mild (weakly acidic or basic) or neutral reaction conditions, to address the issue of functional-group incompatibilities. Although these processes constitute a great advance compared with classical C–C bond-forming methods, the most commonly used catalysts are based on the costly element palladium². Researchers have therefore begun to explore cheaper alternatives to palladium for these reactions. Iron, with its high natural abundance and low cost, is a logical choice.

Many iron-catalysed C–C bond-forming reactions have been discovered in the past decade³. Lo *et al.* were inspired by the ability of iron catalysts to generate reactive free-radical intermediates from alkenes (compounds that contain carbon–carbon double bonds), a property that has been known for more than 20 years⁴. Earlier this year, some of the authors of the current paper reported an iron-catalysed C–C bond-forming process that joins two alkenes together through the intermediacy of a radical⁵. Although useful, this reaction was compatible with only a limited range of functional groups. Lo and colleagues

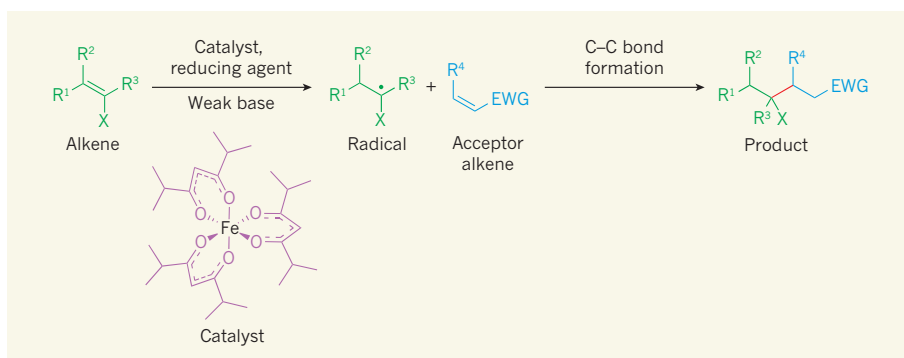


Figure 1 | Iron-catalysed carbon–carbon bond formation with unprecedented functional-group tolerance. Lo *et al.*¹ report an iron catalyst that couples together two alkenes through carbon–carbon (C–C) bond formation. In the presence of a weak base, the catalyst engages a reducing agent to form a species (not shown) that generates a radical from an alkene. This radical adds to an acceptor alkene that incorporates an electron-withdrawing group (EWG), forming a new C–C bond (shown in red in the product). The functional-group tolerance of the reaction derives from the large ligands (purple) bound to the iron atom of the catalyst. R¹–R⁴ represent different carbon-based groups; X represents functional groups containing atoms such as oxygen, nitrogen, sulphur, boron, silicon, fluorine, chlorine, bromine and iodine. The dot on the radical represents a single unpaired electron; broken lines in the catalyst indicate delocalized bonds.

thus set out to develop an improved method that would exhibit broad functional-group tolerance.

The researchers hypothesized that the restricted functional-group tolerance of their original method was caused by the small ligand molecules bound to the iron atom. Accordingly, they prepared and evaluated several iron catalysts that have large ligands. This revealed that a catalyst bearing three bulky diisobutylmethane ligands effectively mediated the formation of a C–C bond between two alkenes in the presence of a reducing agent (Fig. 1). Further investigation established the beneficial effect of a weakly basic additive (disodium phosphate), although its specific role in the reaction is unclear.

Because radical-mediated reactions proceed under mild conditions and typically involve uncharged intermediates, their functional-group tolerance generally exceeds that exhibited by other types of organic reaction, which often involve harsh conditions and charged intermediates^{6,7}. But Lo and co-workers’ reaction exhibits unprecedented functional-group tolerance, even for a process involving radicals. Specifically, atoms such as sulphur, boron, chlorine, bromine and iodine can remain

attached to the radical-bearing carbon atom, emerging unscathed at the end of the reaction; bonds between carbon and these atoms are frequently cleaved during radical processes. The new reaction therefore allows compounds with functional groups containing these atoms to be accessed in a direct and straightforward manner.

Another shortcoming of many C–C bond-forming reactions is the need to rigorously exclude air and moisture from them. Reactions that cannot tolerate the presence of oxygen or water require cumbersome procedures that are difficult to reproduce with precision. Because the new iron-catalysed alkene coupling proceeds efficiently in the presence of water and air, it is conducted using a simple, user-friendly protocol. As a result, anyone with basic training in organic synthesis should be able to successfully perform this reaction and generate reproducible results.

All organic reactions have limitations, and Lo and colleagues’ alkene coupling is no exception. The substrate scope of the alkene that acts as a radical precursor (the green alkene in Fig. 1) is exceptionally broad, but there are some constraints to the structure of the other ‘acceptor’ alkene. Currently, bulky acceptor

alkenes — those with large groups at the R⁴ position shown in Figure 1 — are not viable coupling partners. Further fine-tuning of the catalyst structure and reaction conditions might uncover a solution to this problem.

By facilitating the linking of two alkenes through carbon–carbon bond formation, Lo and co-workers' reaction will allow the direct generation of valuable, structurally complex organic molecules from simpler precursors.

What is more, the iron catalyst is readily prepared from fairly inexpensive ingredients. This method therefore has the potential to transform the way in which chemists think about constructing complicated molecules. ■

Steven L. Castle is in the Department of Chemistry and Biochemistry, Brigham Young University, Provo, Utah 84602, USA.
e-mail: scastle@chem.byu.edu

1. Lo, J. C., Gui, J., Yabe, Y., Pan, C.-M. & Baran, P. S. *Nature* **516**, 343–348 (2014).
2. de Meijere, A. & Diederich, F. (eds) *Metal-Catalyzed Cross-Coupling Reactions* 2nd edn (Wiley, 2004).
3. Nakamura, E. *et al.* *Org. React.* **83**, Ch. 1, 1–209 (2014).
4. Kato, K. & Mukaiyama, T. *Chem. Lett.* **21**, 1137–1140 (1992).
5. Lo, J. C., Yabe, Y. & Baran, P. S. *J. Am. Chem. Soc.* **136**, 1304–1307 (2014).
6. Rowlands, G. J. *Tetrahedron* **65**, 8603–8655 (2009).
7. Rowlands, G. J. *Tetrahedron* **66**, 1593–1636 (2010).

SYNTHETIC BIOLOGY

Toehold gene switches make big footprints

The development of RNA-based devices called toehold switches that regulate translation might usher in an era in which protein production can be linked to almost any RNA input and provide precise, low-cost diagnostics.

**SIMON AUSLÄNDER
& MARTIN FUSSENEGGER**

A fundamental tool of synthetic biology is a type of genetic device that controls the expression of target genes in a trigger-inducible manner, and so can be used to predictably and robustly program cellular behaviour. The number of such gene switches is growing, and switches have been successfully used in combination with other components, such as enzymes to assemble metabolic pathways that produce biofuels¹ and therapeutic drugs², and in designer cells that have the potential to correct metabolic diseases^{3–5}. But the design of circuits of interconnecting switches is often complicated by the fact that each switch is made of natural components and is sensitive to its own predetermined trigger compound. A strategy that produces compatible gene switches tailored to desired trigger compounds would enable the switches to be easily assembled in combination, increasing the precision and complexity with which cellular behaviour can be programmed. Writing in *Cell*, Green *et al.*⁶ describe a method for generating gene switches that can indeed be tailored to desired RNA inputs.

RNA is gathering momentum as a control device for synthetic biology. RNAs are modular, programmable and versatile. Furthermore, the specific sequence of each RNA dictates which molecules it can interact with and what functions its structure confers. The primary RNA sequence is determined by the sequential arrangement of different nucleotides, and this sequence can be engineered so that it forms secondary RNA structures internally or with complementary DNA or RNA molecules. One such structure is the hairpin loop, which

comprises two base-paired sequences ending in an unpaired loop. Secondary structures can affect the translation of messenger RNA, and so can be exploited to regulate protein production from genes of interest.

Translation of mRNA occurs in a complex molecular machine called the ribosome. The ribosome contains a small and a large subunit, both of which are composed of a mixture of ribosomal RNAs and proteins. In bacteria, mRNAs are recruited to the ribosome through their ribosome-binding site (RBS) — a sequence that binds to the small subunit to initiate translation.

The reversible nature of this binding interaction is exploited by a class of engineered RNA-based gene switches called riboregulators, which contain an 'anti-RBS sequence' that binds to the RBS to form a hairpin loop⁷, thus preventing the mRNA from accessing the ribosome and lowering the rate of translation⁶. The anti-RBS sequence is located in the target mRNA itself, in a region that will not be translated into protein, upstream of the site where translation begins. Riboregulators are switched by a 'trigger sequence' that interacts with and disrupts the hairpin, forming an alternative RNA structure that permits RBS–ribosome binding. Depending on the presence or absence of the trigger RNA, target gene expression can therefore be switched on or off. However, because typical riboregulators must fit into the upstream mRNA region and bind to the RBS, they can be designed for only a limited number of trigger sequences.

Green and colleagues have developed a more diverse type of riboregulator, which they call a toehold switch. Toehold riboregulators are designed to interact with the region around the protein-coding start site of each mRNA instead of the RBS, but are not complementary to the start site itself (Fig. 1). Furthermore,

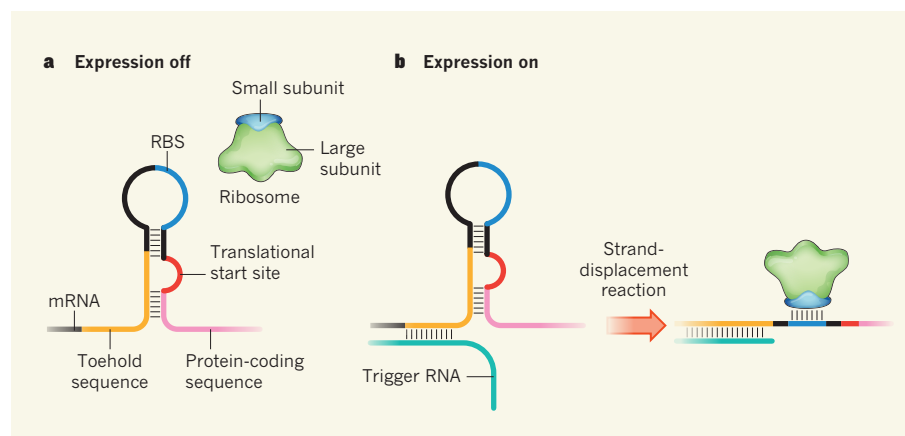


Figure 1 | The design of toehold switches. **a**, Green *et al.*⁶ have designed an RNA-based device, called a toehold switch, that can regulate translation of bacterial messenger RNA in response to the presence or absence of any desired 'trigger' RNA. Toehold switches are located upstream of the site at which translation begins. The switch has an exposed single-stranded region called the toehold sequence that is designed to be complementary to the trigger RNA. To be translated, bacterial mRNA must bind to the small ribosomal subunit through a ribosome-binding site (RBS), but, if the trigger RNA is absent, the presence of the toehold switch causes the formation of a hairpin structure that blocks RBS–ribosome binding, thereby preventing translation. **b**, The presence of the trigger RNA causes a strand-displacement reaction that breaks up the hairpin structure, exposes the RBS to the ribosome and induces translation.

each has an exposed, single-stranded 'toehold sequence' adjacent to the mRNA-binding sequence, which facilitates binding of the trigger RNA to the riboregulator. This design strategy enables riboregulators to be programmed at will.

The authors demonstrated the potential of toehold riboregulators by independently controlling 12 toehold switches inside one cell. They experimentally and computationally characterized their first-generation library of 168 switches to identify specific parameters that are crucial for the proper performance of switches. These parameters enabled the computer-aided design of toehold switches with predictable performance, which was validated for 13 second-generation switches.

These second-generation devices modulated translation extremely efficiently — protein production was up to 650-fold higher when the switch was on than when it was off. This performance is unmatched for other RNA-based switches, and is typically reached only by devices that exert control at the transcriptional level. Demonstrating the versatility of their devices, Green *et al.* produced switches that detect and report on the presence of endogenous RNA sequences, and programmed cellular behaviour using synthetic trigger RNAs.

Might the potential flexibility of toehold switches be exploited in diagnostics? A follow-up report⁸ examined the diagnostic capability of a toehold switch in which the trigger was Ebola virus RNA, and the mRNA under control encoded an enzymatic 'reporter' protein. The switch was freeze dried in cell-free extracts and stored on paper discs. These paper-based switches could reliably detect the Ebola virus RNA with great sensitivity. Furthermore, the switches worked even after long-term storage at ambient temperature. Although designed for use in bacteria, paper-based toehold switches also worked in combination with mammalian cell extracts as protein-based biosensors that quantified blood glucose levels. In the future, paper-based diagnostics might also be used to detect when RNA molecules such as microRNAs are expressed in patterns that are hallmarks of cancer or metabolic disorders⁹.

Although diagnosis is fundamental to any preventive care strategy, therapeutics are also vital, and future treatment strategies could combine the two. Synthetic gene networks that operate inside designer cell implants can monitor, process and score molecular indicators of disease, and can also coordinate the production of protein-based therapies within the engineered cell. Designer networks have been used for the treatment of gouty arthritis³, obesity⁴ and diabetes⁵ in animal models. So far, therapeutic gene networks have used natural sensor components that might be compatible with the human physiological range, but the design of tailor-made biosensors for

specific molecular indicators of disease remains challenging. Toehold switches may be a good starting point to design biosensors specific for any disease-relevant compound — first for microRNAs⁹, and eventually for mutated mRNA sequences. The integration of synthetically engineered biosensors into synthetic gene networks that diagnose and treat disease could dramatically shape cell-based treatment strategies in this century. ■

Simon Ausländer and Martin Fussenegger
are in the Department of Biosystems
Science and Engineering, ETH Zurich,
4058 Basel, Switzerland.

MALARIA

How vector mosquitoes beat the heat

Intensive longitudinal sampling of malaria mosquitoes in the African semi-desert reveals that three morphologically indistinguishable species have distinctive strategies for surviving the dry season. SEE LETTER P.387

NORA J. BESANSKY

The scale-up of interventions against malaria in the past decade has reduced the global death rate of this disease by an impressive 42%. However, more than 600,000 malaria-related deaths still occur each year¹ — 90% of them in sub-Saharan Africa — meaning that malaria remains one of the most significant sources of infectious-disease mortality. Africa has long been recognized as a crucible for malaria-control efforts, owing to its particular blend of widespread and dominant mosquito species that transmit malaria. One of the great mysteries of malariology has been how these vector populations survive the dry season, when there is little water in which the mosquitoes can lay their eggs. In this issue, Dao *et al.*² (page 387) report that they have solved this mystery, but the answer is surprisingly complex, like the vectors themselves.

Three closely related sibling mosquito species belonging to the *Anopheles gambiae* complex are among the most efficient vectors of malaria³ (there are at least seven species in the complex, collectively referred to as *A. gambiae sensu lato* (*s.l.*)). This status is owed to their strong association with humans and their success at exploiting a variety of ecological conditions across tropical Africa, from humid rainforests to the fringes of the Sahara Desert, as long as humans are nearby. However, there is an Achilles heel in the relay of malaria parasites between these vectors and humans — all mosquitoes have an obligate

e-mail: fussenegger@bsse.ethz.ch

1. Peralta-Yahya, P. P., Zhang, F., del Cardayre, S. B. & Keasling, J. D. *Nature* **488**, 320–328 (2012).
2. Paddon, C. J. *et al.* *Nature* **496**, 528–532 (2013).
3. Kemmer, C. *et al.* *Nature Biotechnol.* **28**, 355–360 (2010).
4. Rössger, K., Charpin-El-Hamri, G. & Fussenegger, M. *Nature Commun.* **4**, 2825; <http://dx.doi.org/10.1038/ncomms3825> (2013).
5. Ausländer, S. *et al.* *Mol. Cell* **55**, 397–408 (2014).
6. Green, A. A., Silver, P. A., Collins, J. J. & Yin, P. *Cell* **159**, 925–939 (2014).
7. Ausländer, S. *et al.* *Nature Methods* **11**, 1154–1160 (2014).
8. Pardee, K. *et al.* *Cell* **159**, 940–954 (2014).
9. Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. & Benenson, Y. *Science* **333**, 1307–1311 (2011).

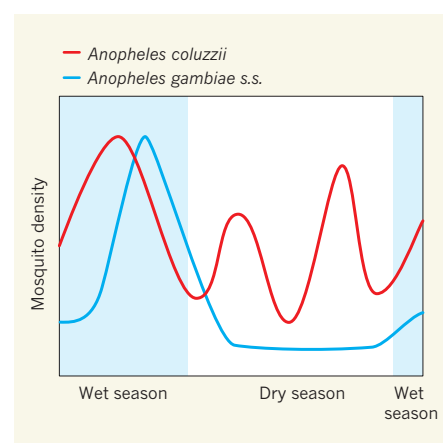


Figure 1 | Species-specific population dynamics. Dao *et al.*² find that average population densities of *Anopheles coluzzii* and *Anopheles gambiae sensu stricto* (*s.s.*) mosquitoes fluctuate seasonally in predictable but distinct patterns. In the wet season, when mosquito breeding sites are abundant and the climate is favourable, densities of both species are high, although *A. coluzzii* achieves its peak population density substantially earlier than *A. gambiae s.s.* does. In the dry season, the *A. gambiae s.s.* population disappears and is not found again until the next wet season, with a slow increase in population density that lags behind that of *A. coluzzii*. By contrast, the *A. coluzzii* population remains in the area during the dry season, but cannot be sampled while the insects are hidden in unknown shelters, leading to apparent troughs. Their emergence from those shelters for two short periods during the dry season is reflected by two peaks in the data.

aquatic immature stage, and in the absence of water, they cannot breed. During the long dry season of the African savannas and the Sahel region, the rains cease for months, surface water evaporates, humidity plummets and temperatures soar. As long as there is no permanent surface water from reservoirs or rivers nearby, malaria transmission becomes undetectable and the local vector mosquitoes also disappear, only to return again with the rains.

Understanding malaria-vector ecology during the dry season, when populations have reached their lowest point, has great strategic significance because deploying mosquito control specifically at those times and places can have the greatest impact. There are two main possibilities for what happens to the mosquitoes during the dry season: long-distance migration to and from refugia where water persists; or stasis, in which the vectors enter a state of dormancy (referred to as aestivation or summer diapause³) that allows them to safely ride out the dry season *in situ*, hidden deep inside (unknown) shelters. Yet finding the disappeared mosquitoes is even harder than it sounds.

In fact, Dao *et al.* did not solve the mystery directly, by physically locating mosquitoes in hiding places or capturing them in the act of long-distance migration, although such efforts are under way³. Instead, their detective work was indirect, using detailed analyses of mosquito population dynamics over time. Although researchers have adopted conceptually similar approaches in the past, the insights that emerge from Dao and colleagues' data were made possible by a sampling effort that is unprecedented both in its detail, allowing the detection of short-lived phenomena, and in its duration, allowing true seasonal patterns to be distinguished from one-off events.

Based in the Sahelian village of Thierola in Mali, the researchers collected mosquitoes from around 120 houses for 2 weeks of every month for 5 years, yielding about 40,000 *A. gambiae* *s.l.* samples. From time-series analysis of the combined data from all three species, the authors inferred a statistically significant repeating seasonal pattern that was unexpectedly complex. They observed the predicted wet-season peak and mid-dry-season trough in vector density, but this was followed by a surprising rise in density in the late dry season, before another low as the dry season ended.

To make biological sense of these data, Dao *et al.* recognized the importance of splitting *A. gambiae* *s.l.* into the three genetically defined units found simultaneously in Thierola: *A. gambiae sensu stricto* (*s.s.*), *Anopheles coluzzii* and *Anopheles arabiensis*. Mosquitoes from the three groups are very closely related and cannot be physically distinguished at any stage in their development. All three hybridize occasionally in nature, but the first two — only recently named as species⁶ and

not universally recognized as such — diverged evolutionarily much more recently than other species in the complex.

Despite the relative youth and morphological homogeneity of this species complex, the fact that the species radiations were accompanied by, if not promoted by, differential adaptations to environmental heterogeneities⁷ makes it unlikely that its members would respond uniformly to a common physiological stress. Notwithstanding this expectation, it is striking that, when Dao and colleagues partitioned the data by species, the two closest relatives (*A. coluzzii* and *A. gambiae* *s.s.*) showed the most distinct population dynamics (Fig. 1). The authors also found that the population density of *A. gambiae* *s.s.* follows a relatively simple pattern of peak abundance in the wet season and a trough throughout the dry season. By contrast, although the density of *A. coluzzii* also peaks in the wet season, the onset of population growth precedes that of *A. gambiae* *s.s.* by two months and, far from disappearing in the dry season, two peaks in population density are consistently observed, despite the absence of rain.

Dao *et al.* make the case that these data best fit a model in which *A. coluzzii* persists locally in a form of diapause and emerges from hiding for two short periods. The cues that provoke this emergence are unknown, but could include abiotic factors, such as increases in humidity or temperature, and biotic factors, such as the need to replenish nutritional reserves — for example, by blood feeding without egg maturation, known as gonotrophic dissociation⁴. By contrast, it seems that *A. gambiae* *s.s.* disappears and, when the rains resume, more slowly recolonizes the area from refugia hundreds of kilometres distant.

Although the population dynamics of *A. arabiensis* were not statistically different from those of *A. gambiae* *s.s.*, small numbers of *A. arabiensis* were collected each dry season, suggesting that at least a fraction of the population remains in place. Whether this implies that the species uses a mixed strategy of diapause and long-distance migration, as the authors propose, or whether there is some other explanation (such as a different type or greater depth of diapause) will require further investigation.

Final proof for these hypotheses will have to come from catching the mosquitoes in the act. Nevertheless, there is now strong evidence that *A. coluzzii* overcomes the stress of the dry season through local diapause, a strategy that ensures its rapid population expansion at the earliest stages of the rainy season and thereby amplifies disease transmission. The long-distance migration proposed for *A. gambiae* *s.s.* will also influence the dynamics of disease transmission and vector control, because both processes determine the ability of vector populations to expand their range and invade distant regions. Unfortunately, we know



50 Years Ago

Dr. H. J. Kingsley and Dr. J. E. A. David of Bulawayo have described ... the case of a girl aged 22 months ... She appeared to be completely insensitive to pain ... She was admitted to hospital for investigation and was noticed to have periods of blankness which were thought to be some type of petit mal. Many investigations were made and all results were normal. While she was in hospital her sensitivity to pain was tested and it was found that the child was insensitive to pain almost all over the trunk, limbs and face, and a sterile hypodermic needle could be stuck through the skin to the subcutaneous tissues without any flinching ... Confusion exists in the literature about congenital absence of pain ... Dr. Walter B. Shelley of Philadelphia thinks that these cases are not as rare as is supposed and that there are people who experience coronary thrombosis or a perforating appendicitis, or have babies, without pain. Apparently, where pain is absent, itching is also absent.

From *Nature* 19 December 1964

100 Years Ago

Physics of the Household. By Prof. C. J. Lynde — The author of this book is professor of physics in the Macdonald College, an affiliated college of the McGill University, Montreal, where a school of household science is one of the branches of the institution, and it is for students of household science that the book is written. It presents the subject of physics in close relation to its domestic applications, and abounds in illustrations and examples of household appliances and processes. It should be of great use to science teachers, especially those who have to teach girls.

From *Nature* 17 December 1914

almost nothing about the environmental cues that prompt these processes, the mechanisms responsible for them or even how generalizable these findings are to mosquito populations elsewhere in tropical Africa. Dao and colleagues' work highlights the urgent need for more field studies to answer these fundamental questions. ■

Nora J. Besansky is at the *Eck Institute for Global Health and Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556-0369, USA.*
e-mail: nbesansk@nd.edu

1. www.who.int/malaria/publications/world_malaria_report_2013/en/
2. Dao, A. *et al. Nature* **516**, 387–390 (2014).

3. Coluzzi, M. *Bull. WHO* **62** (suppl.), 107–113 (1984).
4. Denlinger, D. L. & Armbruster, P. A. *Annu. Rev. Entomol.* **59**, 73–93 (2014).
5. Sohn, E. *Nature* **511**, 144–146 (2014).
6. Coetzee, M. *et al. Zootaxa* **3619**, 246–274 (2013).
7. Powell, J. R., Petrarca, V., della Torre, A., Caccone, A. & Coluzzi, M. *Parassitologia* **41**, 101–113 (1999).

This article was published online on 26 November 2014.

CONSERVATION

Mind the gaps

New analysis reveals the conservation gains that could be achieved by expanding the global network of protected areas — but also how this may be undermined by land-use change and a lack of international coordination. SEE LETTER P.383

THOMAS M. BROOKS

Humanity's best tool for safeguarding nature is the establishment of protected areas¹. Such areas currently cover more than 15% of the terrestrial realm and thus are one of the most extensive uses of the world's land². The Aichi Biodiversity Target 11 — set in 2010 under the Convention on Biological Diversity, a multilateral treaty with 194 parties — states that this proportion should be increased to 17% by 2020. Progress towards Target 11 was a key focus when the international protected-areas community convened in Sydney, Australia, in November 2014 for the sixth World Parks Congress. An analysis by Montesino Pouzols *et al.*³, published online during the congress and now on page 383 of this issue, has direct relevance to these ongoing discussions on protected-area policy and practice.

Understanding the performance of protected areas requires clarity on the extent to which biodiversity is represented in them. Gap analysis — an approach designed to assess how well existing protected areas meet conservation goals — was first applied globally as a contribution to the fifth World Parks Congress, held in 2003 (ref. 4). Montesino Pouzols *et al.* have now taken the method to a new level of sophistication, applying state-of-the-art analytical techniques to massive data sets on biodiversity and protected areas. They find that effective delivery of Target 11, despite the small increase in land proportion covered, could triple current levels of protection of terrestrial vertebrate species and ecological regions (sufficient data are not yet available for an equivalent assessment of invertebrate, plant or fungal species, or of freshwater or marine biomes).

Although this is a crucial finding for

global-level policy, it does not directly inform where to fill the gaps in the current protected-area network on the ground. This is because Montesino Pouzols and colleagues use a spatial resolution of 0.2 degrees (equating to squares of about 20 kilometres on each side at the Equator), which, although finer than the resolution of the biodiversity data they analyse, is coarse compared to the resolution of actual protected-area boundaries. However, the authors did run a sensitivity analysis that revealed similar results at spatial resolutions as fine as 1/60 degrees, demonstrating that their aggregate results are robust.

The authors also validate their findings using high-resolution assessments of key biodiversity areas identified through existing national-level analyses for the Philippines, Myanmar and Madagascar. These are sites that significantly contribute to global biodiversity persistence, documented using standard criteria such as those used by the conservation partnership BirdLife International for the identification of Important Bird and Biodiversity Areas or by the Alliance for Zero Extinction for identifying sites that are the single remaining home of one or more highly threatened species⁵. The consultation currently under way to finalize the key biodiversity-area standard was discussed at length at the Sydney World Parks Congress. The authors' validation in turn suggests the way forward to fill protected-area gaps, by using key biodiversity areas as critical inputs for systematic resource allocation and conservation planning. Protected-area coverage of key biodiversity areas is already used⁶ as a marker of progress towards Target 11 and has been proposed⁷ as indicator 87 towards the United Nations' putative Sustainable Development Goal 15.

However, Montesino Pouzols and co-workers also reveal dark clouds on the protected-area horizon. Their modelling shows that projected scenarios for land-use changes that degrade or eliminate habitats and thus preclude protection will foreclose options for efficient achievement of Target 11. What might bring light into this future shadow? Perhaps most important will be to recognize and document protected-area governance and "other effective area-based conservation measures" — which are often currently focused on governmental roles — as including private protected areas and indigenous and community conserved areas⁸. Many key biodiversity areas beyond the current protected-area network are already



Figure 1 | Community protection. The Reserva Natural El Pangán in Colombia is recognized as an Important Bird and Biodiversity Area. Although not designated as a governmental protected area, the Colombian non-governmental organization ProAves works with the local community to safeguard the site, with funding from the Critical Ecosystem Partnership Fund and other international sources — an example of biodiversity protection beyond that afforded by government-designated sites.

PROAVES COLOMBIA

managed as de facto protected areas (Fig. 1). With better documentation and recognition of their contributions, it is probable that more such sites would be designated. Two entire streams of discussion in Sydney addressed this issue, and Montesino Pouzols and colleagues' finding underscores the urgency of filling protected-area gaps, whether through governmental or non-governmental mechanisms. The longer the delay in implementation, the more challenging and expensive such efforts will become.

The second challenge revealed by the authors has been little appreciated until now, but they find that it is more than twice as severe as impending habitat conversion. By comparing their global gap analysis with how actions towards Target 11 might be implemented independently by each of the world's countries, they show that uncoordinated national protected-area establishment would compromise efficiency by one-third or more. This is because biodiversity is distributed so unevenly around the planet. Thus, disproportionate responsibility lies with the more biodiverse, and often less developed, tropical nations, whereas the rich northern countries are those with not only the least biodiversity, but also the highest local costs⁹.

This discrepancy could be resolved by adopting the concept of common but differentiated responsibility, placing greater attention on international support to the US\$76.1 billion of annual financing that has been estimated as necessary¹⁰ to achieve Target 11. Rather than designating new protected areas within their borders, developed countries could make greater contributions to conservation targets by supporting the establishment of protected areas in tropical countries, where the biodiversity benefit per dollar would be substantially higher. The largest fund to facilitate such programmes is the Global Environment Facility; its work is complemented by programmes such as the Critical Ecosystem Partnership Fund. At the Sydney congress, both institutions reiterated their commitments to deliver incremental global biodiversity benefits by channelling and coordinating such resources.

Current indicators towards the Aichi Biodiversity Targets focus on nations' actions within their own borders⁶. Thus, incentivizing coordinated international funding will require the establishment of parallel indicators of the conservation benefits of investment by each country. Such developments are essential, and urgently needed, if we are to close the gaps in implementing Target 11, by staving off the challenges of future threats and of national self-interest identified so incisively by Montesino Pouzols and colleagues. ■

Thomas M. Brooks is at the International Union for Conservation of Nature, Gland 1196, Switzerland.
e-mail: thomas.brooks@iucn.org

1. Watson, J. E. M., Dudley, N., Segan, D. B. & Hockings, M. *Nature* **515**, 67–73 (2014).
2. Juffe-Bignoli, D. et al. *Protected Planet Report 2014* (UNEP-WCMC, 2014); available at go.nature.com/wi5t3x
3. Montesino Pouzols, F. et al. *Nature* **516**, 383–386 (2014).
4. Rodrigues, A. S. et al. *Nature* **428**, 640–643 (2004).
5. Butchart, S. H. M. et al. *PLoS ONE* **7**, e32529 (2012).
6. Tittensor, D. P. et al. *Science* **346**, 241–244 (2014).
7. Sachs, J. D. & Schmidt-Traub, G. (eds) *Indicators and a Monitoring Framework for Sustainable Development Goals* (Sustainable Development Solutions Network, 2014); available at go.nature.com/xkk1bl
8. Borri-Feyerabend, G. et al. *Governance of Protected Areas* (IUCN, 2014); available at go.nature.com/3kjqnj
9. Balmford, A., Gaston, K. J., Blyth, S., James, A. & Kapos, V. *Proc. Natl Acad. Sci. USA* **100**, 1046–1050 (2003).
10. McCarthy, D. P. et al. *Science* **338**, 946–949 (2012).

MATERIALS SCIENCE

Two steps for a magnetoelectric switch

Magnetoelectric materials allow magnetism to be controlled by an electric field. The discovery of an indirect path for switching electrical polarization in one such material brings this idea close to practical use. SEE LETTER P.370

KATHRIN DÖRR & ANDREAS HERKLOTZ

Devices called spin valves have become crucial for magnetic sensing and data storage^{1,2}. They consist of two layers of ferromagnetic materials (which exhibit the familiar form of magnetism found in iron bar magnets), and have low electrical resistance when the direction of magnetization of the layers is the same, but high resistance when the magnetizations are antiparallel. Such devices can be switched between these states using a magnetic field or a large spin-polarized current — an electrical current in which the majority of electrons have the same spin orientation. On page 370 of this issue, Heron et al.³ report a method to control spin-valve states at room temperature using an electric field. Their approach consumes much less energy than using spin-polarized current, and

opens up opportunities for miniaturizing spin valves that are not possible for magnetically controlled devices.

In magnetoelectric materials, electric dipoles associated with atoms, ions or molecules can be ordered using a magnetic field, and magnetic moments can be ordered using an electric field. Magnetoelectricity is an extremely rare property, most likely to occur in 'multiferroic' materials, in which spontaneous magnetic order and ferroelectric behaviour coexist⁴; ferroelectricity is spontaneous polarization of electric dipoles that can be reversed using an electric field. Bismuth ferrite (BiFeO_3) is the only material known to show robust multiferroicity up to high temperatures: it is ferroelectric below 1,100 kelvin and also antiferromagnetic below 640 K (that is, the magnetic moments of its iron atoms alternate in direction below this temperature, yielding a

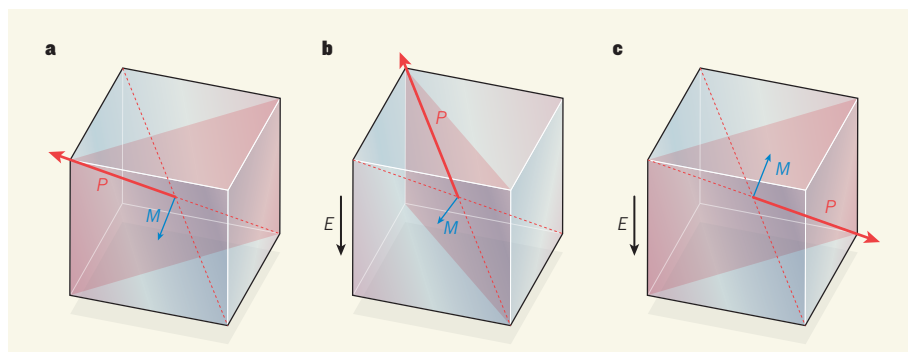


Figure 1 | Indirect reversal of magnetization using an electric field. **a**, In this depiction of the unit cell of strained bismuth ferrite, spontaneous polarization of electric dipoles (known as ferroelectric polarization, P) is constrained to one of the space diagonals of the unit cell. Weak magnetization (M) occurs at right angles to P . **b, c**, Heron et al.³ report that an electric field (E) reverses the direction of P in a two-step process. P first changes orientation by 71° , aligning with a second space diagonal (**b**), and then undergoes a further shift of 109° , returning to the first space diagonal (**c**). Because M is coupled to P , it undergoes an analogous two-step reversal.

net magnetization of zero). However, a small magnetization arises from canting (tilting) of the anti-aligned magnetic moments.

The weak magnetization of bismuth ferrite can be amplified through strong interfacial magnetic coupling to a ferromagnet^{5,6}. So, if the weak magnetization could be coupled to the material's ferroelectric polarization, then the direction of magnetization of an adjacent ferromagnetic layer could be altered using an electric field. Unfortunately, direct reversal of ferroelectric polarization in bismuth ferrite was expected to leave the orientation of the canted magnetic moment unchanged⁷, which means that magnetoelectrically driven magnetization reversal would be impossible.

Enter Heron and colleagues, who find that ferroelectric-polarization reversal in a strained bismuth ferrite film follows an indirect pathway. When grown as an elastically strained layer on a substrate of DyScO₃ (Dy is dysprosium; Sc is scandium) covered with an electrode of strontium ruthenate (SrRuO₃), bismuth ferrite forms stripe-like ferroelectric domains containing stable polarization of electric dipoles, in which the orientations of the dipoles are constrained to two of the four space diagonals of the material's pseudocubic unit cell. The authors studied local ferroelectric switching in the electric field of the tip of a scanning force microscope at the surface of the bismuth ferrite, obtaining data at microsecond resolution during repeated scans under an applied direct-current voltage. They observed an overall full reversal of polarization (a 180° switch) nearly everywhere, but this was reached through an intermediate switch by 71° or 109° via the other 'active' space diagonal of the unit cell (Fig. 1). The researchers' computational modelling confirmed that the energy barrier to this two-step switching is smaller than for direct 180° switching, and predicted that the indirect switching path reverses the canted magnetic moment.

The authors next deposited a spin valve consisting of two layers of a ferromagnet (Co₉Fe_{0.1}) separated by a copper layer onto an underlayer of bismuth ferrite. They found that the device exhibited about the same values of high and low electrical resistance in a field — either in an in-plane magnetic field or in an electric field at the vertical to the bismuth ferrite layer — when the field was cycled between positive and negative values. This result requires the magnetization of the Co₉Fe_{0.1} to be aligned with the canted magnetic moment in the adjacent bismuth ferrite film⁶. The authors used a technique called X-ray magnetic circular dichroism photo-emission electron microscopy to show the local magnetization reversal that occurs in Co₉Fe_{0.1} after voltage pulses are applied to the bismuth ferrite underlayer, demonstrating that magnetic coupling at the interface is strong enough to induce the switching.

The crucial message from Heron and colleagues' work is that the switching path

matters: the final orientation of coupled ferroic order parameters (such as spontaneous magnetization, electric polarization and elastic strain) during a switching process depends not only on the initial orientation of the parameters and the direction of applied fields, but also on intermediate steps during switching. These steps are governed by kinetic barriers that affect the dynamics of the switching process, and can be assessed using theory and controlled by elastic strains in films. The authors' findings open up the prospect of multi-step switching processes that access unexpected states of multiferroic materials other than bismuth ferrite. This in turn might enable alternative strategies for engineering magnetoelectric switching. The study also confirms that the weak magnetism common to many multiferroics can be amplified by magnetic coupling to a conventional ferromagnetic metal⁵.

Techniques that allow high-quality interfaces between ferromagnetic metals and multiferroics to be prepared have enabled experiments that advance our understanding of magnetic coupling at such interfaces⁶. However, these interfaces are unstable in large, cycling electric fields, and this issue must be resolved for practical applications. The same problem afflicts interfaces between ferroelectric materials and metal electrodes, but has been partly solved following long-standing research efforts⁸. Nevertheless, the stability of ferroelectric-metal interfaces is still a major factor in the lifetimes of devices that use such interfaces.

INFLUENZA

An RNA-synthesizing machine

Crystal structures of the complete RNA polymerases from influenza A and B viruses provide insight into how these enzymes initiate RNA synthesis, and reveal targets for antiviral drug design. SEE ARTICLES P.355 & P.361

ROBERT M. KRUG

Influenza A and B viruses cause a highly contagious disease in humans that results in approximately 250,000 to 500,000 deaths worldwide each year¹. In addition, influenza A viruses are responsible for periodic human pandemics that can have substantially higher mortality rates; the most severe pandemic, in 1918, caused around 40 million deaths². The primary defence against influenza infections has been vaccination, but antiviral drugs also play a key part, for example in the elderly, who are not well protected by vaccines. Circulating influenza viruses have developed

The walls of newly formed ferroelectric domains cross from one side to the other of the stripe domains observed by Heron *et al.* during switching. So, if the stripe domains can be engineered to be narrow (of the order of 100 nanometres wide), this would help the fast operation of multiferroic spintronic devices, such as spin valves, by reducing the crossing time to a few nanoseconds in a sufficiently large electric field. A considerable amount of work may be needed to develop fast magnetoelectric switching and to stabilize multiferroic-to-ferromagnet interfaces. Even so, the electrical control of ferromagnetism demonstrated by Heron and co-workers is a decisive step towards the realization of multiferroic spintronic devices. ■

Kathrin Dörr is at the Institute for Physics, Martin Luther University Halle-Wittenberg, 06099 Halle, Germany. **Andreas Herklotz** is in the Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA. e-mail: kathrin.doerr@physik.uni-halle.de

1. Grünberg, P. A. *Rev. Mod. Phys.* **80**, 1531–1540 (2008).
2. Fert, A. *Rev. Mod. Phys.* **80**, 1517–1530 (2008).
3. Heron, J. T. *et al. Nature* **516**, 370–373 (2014).
4. Spaldin, N. A. & Fiebig, M. *Science* **309**, 391–392 (2005).
5. Heron, J. T. *et al. Phys. Rev. Lett.* **107**, 217202 (2011).
6. Trassin, M. *et al. Phys. Rev. B* **87**, 134426 (2013).
7. Ederer, C. & Spaldin, N. A. *Phys. Rev. B* **71**, 060401(R) (2005).
8. Boyn, S. *et al. Appl. Phys. Lett.* **104**, 052909 (2014).

resistance to several of the available antivirals^{3,4}, highlighting the need to develop new drugs. The viral RNA polymerase — the enzyme that catalyses the synthesis of virus-specific RNAs in infected cells — is an attractive target for antiviral drug development, but progress has been hampered by the absence of a three-dimensional structure of the enzyme. Two papers from the Cusack research group published in this issue (Pflug *et al.*⁵ on page 355 and Reich *et al.*⁶ on page 361) now report the first structures of the viral polymerases of influenza A and B, respectively.

Influenza A and B viral polymerases are each composed of three proteins: PB1, PB2

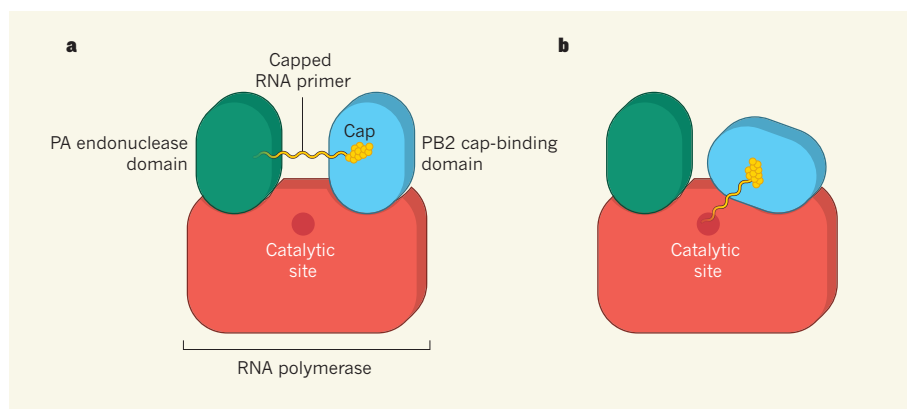


Figure 1 | RNA transcription by influenza polymerase. The influenza A and B RNA polymerases are made up of three viral proteins: PB1, PB2 and PA. The complete structures presented by Pflug *et al.*⁵ and Reich *et al.*⁶ show the location of key elements of these protein complexes. **a**, Synthesis of viral messenger RNA is initiated by a process called cap-snatching, in which a short RNA primer, containing a cap structure, is cleaved from a cellular pre-mRNA (not shown) by the PA endonuclease domain of the viral polymerase. The PB cap-binding domain and the endonuclease domain at the two ends of the U-shaped viral polymerase initially face each other across a channel, and the endonuclease cleaves the pre-mRNA to produce the capped RNA primer. **b**, The cap-binding domain then moves the capped primer away from the endonuclease domain and directs it down into the RNA-synthesizing catalytic site.

and PA (ref. 7). The tripartite polymerase is activated by RNA sequences found at the two extremities of the viral RNA, which interact with each other to form a partially double-stranded structure referred to as the promoter for viral RNA synthesis. The new three-dimensional structures contain the viral polymerase in association with this RNA promoter.

The viral polymerases catalyse two types of RNA synthesis: transcription to produce messenger RNA; and RNA replication to produce templates for the production of viral RNA. These two processes are initiated by different mechanisms. The synthesis of viral mRNA is initiated by a process called cap-snatching, which was discovered more than 30 years ago⁸. The viral polymerase binds to a chemical structure, the cap, located at the ends of the precursors of cellular mRNAs (pre-mRNAs), and the polymerase then uses its endonuclease (nucleic-acid cleaving) activity to cleave the pre-mRNAs at a position 10–15 nucleotides downstream from the cap. The resulting cap-containing fragment acts as a primer to initiate viral mRNA synthesis. By contrast, the RNA-replication reaction is initiated without a primer⁷. The viral polymerases analysed in the new papers catalyse the initiation of both transcription and RNA replication.

The structures of the influenza A and B RNA polymerases show that the three protein subunits make multiple complex interactions with each other, and that all three subunits participate in most of the key polymerase functions. The polymerase forms a U-shaped structure, with the cap-binding domain (which is part of the PB2 protein) at the top of one of the arms of this structure and the endonuclease domain (which is located in the amino-terminal region of the PA protein) at the top of the other arm. All

three polymerase subunits are involved in positioning the endonuclease domain. In the influenza A polymerase structure, the cap-binding and endonuclease domains face each other across a channel whose breadth corresponds to the length of the cap-containing fragment that is produced (Fig. 1).

Reich *et al.* determined two influenza B polymerase structures. On the basis of their observation that the cap-binding domain of one of these crystal structures is rotated by 70° compared to the other B polymerase structure and Pflug and colleagues' influenza A polymerase structure, the authors propose that this domain moves the capped primer away from the endonuclease domain and directs it down into the RNA-synthesizing catalytic site of the polymerase, which is in the PB1 protein (Fig. 1). The N-terminal endonuclease-containing domain of PA is on the opposite side of the polymerase to the larger carboxy-terminal PA domain. The two PA domains are connected by a PA linker region that wraps around the external face of the PB1 protein. All three polymerase protein subunits participate in the binding of the two strands of the RNA promoter that is located close to the PB1 active site.

Reich and colleagues deduce a possible mechanism by which viral RNA replication is initiated in the absence of a primer, by comparing the influenza A polymerase structure with the structures of other viral RNA polymerases, particularly those of hepatitis C and dengue viruses. These two viral polymerases contain a 'priming loop' in which an aromatic amino-acid residue (tyrosine or tryptophan) stabilizes the positioning of the initial ribonucleotide triphosphates that form base pairs with the template RNA in the absence of an RNA primer⁹. The PB1 protein in the

influenza A polymerase contains an analogous putative priming loop (amino acids 641–657) in which histidine (an aromatic amino-acid residue) at position 649 could interact with the initial incoming ribonucleoside triphosphates to facilitate unprimed initiation.

Pflug and colleagues' influenza A polymerase structure includes the region of the PB2 protein that has been implicated in the adaptation of avian influenza A viruses for replication in mammalian hosts¹⁰, and this should provide insight into the mechanisms by which this region functions. Indeed, the polymerase structures will provide the basis for much future work. Previous structures of fragments of the polymerase identified the cap-binding and endonuclease sites^{11–13}, both of which are potential targets for the development of anti-influenza drugs. The new, complete structures provide further targets, including the PB1 active site, the binding sites of the two strands of the viral RNA promoter and numerous sites that are required for essential rearrangements during viral RNA synthesis.

The structures also give us a much better understanding of the mechanisms of influenza virus RNA synthesis — both for transcription and viral RNA replication. Nonetheless, there are considerable gaps in our knowledge of these processes. For example, we do not know how the 3' end of the viral RNA template is relocated to the PB1 active site, nor how the polymerase progresses from initiation of viral RNA synthesis to the elongation of RNA chains. Furthermore, because the viral RNA template in infected cells is coated with multiple viral nucleoprotein molecules along almost its entire length, it is not clear how the viral RNA polymerase accesses and copies such a template. ■

Robert M. Krug is in the Department of Molecular Biosciences, Center for Infectious Disease, Institute of Cellular and Molecular Biology, University of Texas at Austin, Austin, Texas 78712, USA.
e-mail: rkrug@austin.utexas.edu

1. World Health Organization. www.who.int/mediacentre/factsheets/fs211/en/
2. Reid, A. H., Taubenberger, J. K. & Fanning, T. G. *Microbes Infect.* **3**, 81–87 (2001).
3. Deyde, V. M. *et al.* *J. Infect. Dis.* **196**, 249–257 (2007).
4. Bloom, J. D., Gong, L. I. & Baltimore, D. *Science* **328**, 1272–1275 (2010).
5. Pflug, A., Guilligay, D., Reich, S. & Cusack, S. *Nature* **516**, 355–360 (2014).
6. Reich, S. *et al.* *Nature* **516**, 361–366 (2014).
7. Boivin, S., Cusack, S., Ruigrok, R. W. & Hart, D. J. *J. Biol. Chem.* **285**, 28411–28417 (2010).
8. Plotch, S. J., Bouloy, M. & Krug, R. M. *Proc. Natl Acad. Sci. USA* **76**, 1618–1622 (1979).
9. Lescar, J. & Canard, B. *Curr. Opin. Struct. Biol.* **19**, 759–767 (2009).
10. Cauldwell, A. V., Long, J. S., Moncorgé, O. & Barclay, W. S. *J. Gen. Virol.* **95**, 1193–1210 (2014).
11. Dias, A. *et al.* *Nature* **458**, 914–918 (2009).
12. Yuan, P. *et al.* *Nature* **458**, 909–913 (2009).
13. Guilligay, D. *et al.* *Nature Struct. Mol. Biol.* **15**, 500–506 (2008).

2014

EDITORS' CHOICE

Extracts from selected
News & Views articles published
this year.

POPULATION HEALTH

IMMATURITY IN THE GUT MICROBIAL COMMUNITY

Elizabeth K. Costello & David A. Relman
(*Nature* 510, 344–345; 2014)

Physical measurements, such as weight for height, scored relative to a reference population are indispensable tools in the assessment and treatment of undernutrition. But Subramanian *et al.* have charted a different path — one in which the milestones are microbial — for young children living in the Mirpur urban slum of Dhaka, Bangladesh. By surveying the bacterial communities in faecal samples from 50 well-nourished subjects, the authors defined two indicators of gut-microbiota maturation: relative microbiota maturity and a microbiota-for-age Z-score. Compared with healthy children, malnourished children showed significant microbiota immaturity. In the 2–3 months following treatment, the children's microbiota-maturation scores improved; however, after this period, much of this catch-up maturation was lost. The approach presented by the authors could be used to develop standards across the globe, and then to monitor gut colonization during early childhood, as an early-warning system for microbiotas that are falling 'off track'.
Nature 510, 417–421 (2014).

ACCELERATOR PHYSICS

SURF'S UP AT SLAC

Mike Downer & Rafal Zgadzaj (*Nature* 515, 40–41; 2014)

In November 2012, Guinness World Records reported that 120 surfers in Australia rode the same wave simultaneously for more than 5 seconds. "The trick was to get them all to do the same thing at the same time," said group leader Wes Smith. "It was an operation of military-like precision and we finally got there." Now Litos and colleagues, in work at the SLAC National Accelerator Laboratory, have 'got there', too, by surfing half a billion 20-billion-electronvolt electrons on a steep charge-density wave about the size of a marine phytoplankton, travelling through ionized gas (plasma). The wave was driven by a companion electron bunch as it raced at nearly the speed of light through a 30-centimetre-long chamber filled with plasma. Although this inaugural experiment lost about 90% of its 'surfers' along the way, the surviving electrons gained 1.6 billion electronvolts in energy with unparalleled uniformity, maintaining roughly 1% energy spread throughout their wild ride, while sucking away an unprecedented fraction (up to 30%) of the wave's energy. The result might herald a new generation of compact 'plasma afterburners' that could boost the energy of conventional particle accelerators and potentially reduce the skyrocketing cost of high-energy physics machinery.
Nature 515, 92–95 (2014).



SYLVIE DEMOUCY

EARTH SCIENCE

MISSING LINK IN MANTLE DYNAMICS

Greg Hirth (*Nature* 507, 42–43; 2014)

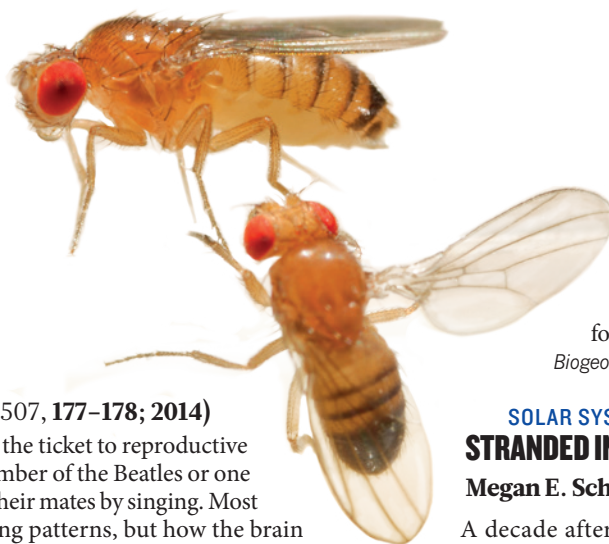
The viscosity of Earth's upper mantle controls a wide range of processes, from the attenuation of seismic waves and the rate of surface deformation after earthquakes to the slow, global-scale flow that is associated with mantle convection and the dynamics of tectonic plates. This viscosity is logically interpreted as being dominated by the physical properties of olivine, the most abundant mineral in Earth's upper mantle, as well as in those of the other terrestrial planets (Mars, Venus and Mercury) and the Moon. Cordier *et al.* report how new techniques to analyse the microstructure of grain boundaries in olivine (pictured) allowed them to discover crystal defects called disclinations in this mineral. This observation is probably a first for geological materials, and has ramifications for our understanding of the processes that control mantle dynamics.
Nature 507, 51–56 (2014).

EVOLUTIONARY DEVELOPMENTAL BIOLOGY

USE IT OR LOSE IT

Bau-lin Huang & Susan Mackem (*Nature* 511, 34–35; 2014)

Adaptive digit loss enables specialized functions such as running or flight, and has repeatedly evolved in parallel. But the developmental mechanisms underlying deviation from the five-digit ground state are unclear, partly owing to the hurdles involved in analysing embryos from animals that are not typically studied in the laboratory. Cooper *et al.* find that, in three-toed jerboas, expanded apoptotic regions encompass the digit I and V precursors, and that camels and horses also use such cell-death mechanisms to reduce digit number. But the cell-death machinery is unaltered in some ungulates. Instead, Cooper *et al.* and Lopez-Rios *et al.* implicate reduced *Ptch1* gene expression as the primary basis for digit loss in cows and pigs, although the extent of digit reduction differs. Further dissecting the diverse mechanisms converging on similar structures will require characterization of the tissue-specific regulation for each candidate gene.
Nature 511, 41–45, 46–51 (2014).

365 DAYS:
the year in science

NEUROSCIENCE

ORDERED RANDOMNESS
IN FLY LOVE SONGSBence P. Ölveczky (*Nature* 507, 177–178; 2014)

Well-crafted love songs can be the ticket to reproductive success, whether you are a member of the Beatles or one of the many animals that woo their mates by singing. Most animals vary their mating-song patterns, but how the brain generates such variability remains a mystery. To address this, Coen *et al.* focus on the song of the male fruit fly. Just as the Beatles made a career of mixing ‘love’, ‘you’, ‘me’, ‘she’ and ‘baby’ in different ways, so male flies switch between ‘sine’ and ‘pulse’ songs. The authors find that the male’s visual experience of the female’s movements shapes his song through neural circuits that control locomotion. In fact, the best predictor of song structure is not the female’s movements, but the singer’s own. The picture that emerges is one in which the male executes a tightly integrated song-and-dance number, inspired by his partner’s movements. This study demonstrates that detailed analysis can distil seemingly complex and unpredictable behavioural patterns into simple rules and sensorimotor transformations.

Nature 507, 233–237 (2014).

BIOGEOSCIENCE

AFRICA’S GREENHOUSE-GAS BUDGET IS IN THE RED

Cheikh Mbow (*Nature* 508, 192–193; 2014)

One of the biggest challenges in curbing climate change is to obtain robust estimates of greenhouse-gas emissions and sequestration. Writing in *Biogeosciences*, Valentini *et al.* have risen to this challenge by providing a full greenhouse-gas assessment for Africa. Until now, scientific opinion has held that Africa could help to reduce emissions or sequester carbon if deforestation of large areas could be avoided, or if tree and forest cover could be increased through sustainable practices, for example agroforestry (pictured) and plantation management. The authors concur that Africa is a small carbon sink on an annual timescale. But, more surprisingly, they find that it may be a net source of radiative forcing — reradiation of heat back towards Earth’s surface by greenhouse-gas molecules — when the greenhouse gases methane and nitrous oxide are included in the annual budgeting. This paper could instigate



JAKE LVELL/ALAMY

a turning point in the design of programmes for counteracting climate change in Africa.

Biogeosciences 11, 381–407 (2014).

SOLAR SYSTEM

STRANDED IN NO-MAN’S-LAND

Megan E. Schwamb (*Nature* 507, 435–436; 2014)

A decade after its discovery, Sedna still remains one of the strangest objects in the Solar System. This remote icy body has a highly eccentric orbit that extends to about 1,000 astronomical units and has a perihelion of 76 AU. Its orbit is well beyond the reach of Neptune, which is located at 30 AU, and is a long way from the edge of the Solar System, where the Oort cloud, the reservoir of long-orbital-period comets, resides at about 10,000 AU. Although other potential candidates have been found, Sedna had remained the solitary confirmed member of a proposed inner Oort cloud beyond 70 AU. Trujillo and Sheppard report the discovery of an object, called 2012 VP₁₁₃, which joins Sedna as the second confirmed member of the inner Oort cloud. The finding solidifies the existence of a population of icy bodies probably ranging in size from a few to a thousand kilometres.

Nature 507, 471–474 (2014).FORUM: Synthetic biology
ENGINEERING EXPLORED(*Nature* 509, 166–167; 2014)

The aim of synthetic biology is to predictably bioengineer organisms that perform beneficial functions. This involves modifying and reassembling biological components. Two views are presented here on the best way to engineer these components so that they reliably generate organisms with desired traits.

RATIONALIZING NATURE

Pamela A. Silver & Jeffrey C. Way

The gene was originally defined as the basic biological unit. But manipulation of DNA has revealed tantalizing levels of modularity that extend to many other cellular regulatory elements. The dream of the rational designer is to understand these modular parts in sufficient detail to be able to assemble them logically, much as an engineer would build a machine for a certain purpose. There are already success stories — for example, assembly of simple genetic circuits that rely on the existence of two stable states in a system.

EVOLVING WITH PURPOSE

Frances H. Arnold & Joseph T. Meyerowitz

Synthetic biologists cannot yet create an enzyme or a biosynthetic pathway that compares favourably with nature’s engineering outputs. The reason is simple: in biology, details matter a lot, and we don’t understand the details. Rational design will not move forward until our understanding of the details of biology has improved dramatically. Luckily, we do not have to wait. Evolution is a time-tested tool for engineering the details, and we can use it in the lab to circumvent our profound ignorance of how sequence encodes function. ■

Functionalized olefin cross-coupling to construct carbon–carbon bonds

Julian C. Lo^{1*}, Jinghan Gui^{1*}, Yuki Yabe¹, Chung-Mao Pan¹ & Phil S. Baran¹

Carbon–carbon (C–C) bonds form the backbone of many important molecules, including polymers, dyes and pharmaceutical agents. The development of new methods to create these essential connections in a rapid and practical fashion has been the focus of numerous organic chemists. This endeavour relies heavily on the ability to form C–C bonds in the presence of sensitive functional groups and congested structural environments. Here we report a chemical transformation that allows the facile construction of highly substituted and uniquely functionalized C–C bonds. Using a simple iron catalyst, an inexpensive silane and a benign solvent under ambient atmosphere, heteroatom-substituted olefins are easily reacted with electron-deficient olefins to create molecular architectures that were previously difficult or impossible to access. More than 60 examples are presented with a wide array of substrates, demonstrating the chemoselectivity and mildness of this simple reaction.

New methods for the construction of C–C bonds have the potential to shift paradigms in retrosynthetic analysis (the strategy used to design syntheses of molecules)¹. Historically, those that have been most successful feature simple experimental procedures, exhibit broad scope and allow access to chemical space previously deemed challenging or inaccessible. A recent exercise in total synthesis drew our attention to radical-based olefin hydrofunctionalizations of the sorts pioneered in refs 2–9. Those illuminating studies led to the invention of a reductive coupling^{10–12} of simple olefins with electron-deficient olefins such as that depicted in Fig. 1a¹³. In that work, an adduct bearing an all-carbon quaternary centre such as **A** could be easily accessed in minutes and in an open flask from olefin **B**, presumably via the intermediacy of radical **A'**. Although a useful and practical method, the compounds it produced could already be obtained from readily accessible functionalized hydrocarbons such as alkyl halides¹⁴, alcohols^{15,16} and carboxylic acids¹⁷ via conventional radical-generating processes.

In contrast, the functionalized hydrocarbons required to access adducts such as **C**, **D** and **E** (Fig. 1a) would either require extensive functional group (FG) manipulations or are unfeasible donors owing to FG incompatibilities and chemoselectivity difficulties arising from the heteroatoms present (**B**, **S** and **I**). By analogy to previous work, if olefins could be used as a surrogate for the intermediate radicals **C'**, **D'** and **E'**, easily accessible compounds such as **F** could be employed directly, avoiding FG manipulations completely.

Development of functionalized olefin cross-coupling

Although this idea is conceptually simple, examining the hypothetical mechanistic pathway revealed numerous obstacles that would need to be addressed, as shown in Fig. 1b. The initiating step, radical formation from the donor olefin **G** by an *in situ*-generated Fe hydride, could be complicated by issues of both regioselectivity and chemoselectivity. Furthermore, depending on the nature of the X substituent, several competing pathways could arise involving the Fe complexes in the catalytic cycle (for example, transmetalation of a C–B bond, desulfurization of a C–S bond, and oxidative addition of a C–I bond). If the first step did occur as intended, the intermediate radical **H** could be prone to premature reduction^{18–20}, trapping with O₂ (ref. 2), or homodimerization.

Provided that **H** undergoes the desired conjugate addition to the electron-deficient olefin coupling partner, the newly generated radical **I** could undergo homodimerization, intramolecular hydrogen atom abstraction or consecutive conjugate additions leading to uncontrollable oligomerization. Formation of **J** from a single-electron reduction of **I** would result in a substantially basic and nucleophilic site that could prove to be incompatible with the X group and its substituents. In order for the reaction to prove successful, the conditions must be mild enough to tolerate both the various intermediate species in the catalytic cycle, as well as the final coupled product **K**.

With these potential difficulties in mind, we used the model system depicted in Fig. 2a, with silyl enol ether **1** serving as the donor and cyclohexenone (**2**) as the acceptor, to develop a functionalized olefin cross-coupling. Application of conditions similar to those previously developed, using Fe(acac)₃ (**4**, acac, acetylacetonate) as a catalyst and PhSiH₃ as a stoichiometric reductant^{3,13}, formed the reductively coupled product **3** in 53% yield based on GC/MS (gas chromatography/mass spectrometry) using an internal standard. Analysis of the side products from the model system and related reactions led to the identification of compounds **14–17** (Fig. 2b). As **16** and **17** presumably arise from pathways where Fe(acac)₃ behaves as a Lewis acid²¹, we hoped to attenuate the Lewis acidity of the catalyst by increasing the amount of steric shielding of the Fe centre. Increasing the size of the substitution on the diene ligands (**5–9**) led to decreased amounts of **16**, with Fe(dibm)₃ (**5**, dibm, diisobutrylmethane)²² providing the best balance between reactivity and steric shielding. Although attempts to alter the electronic structure of the ligand with electron-deficient (**10** and **11**) and electron-rich (**12** and **13**) substituents eliminated reactivity, the addition of Na₂HPO₄ increased the yield of the desired product **3** from 69% to 78% when using Fe(dibm)₃ as the catalyst. The use of about 45 other inorganic and amine bases as additives did not result in increased yields, suggesting that Na₂HPO₄ does not simply serve as a buffering agent. Additionally, Fe(dibm)₃ enabled product formation with donors that were unreactive with Fe(acac)₃ (**18**, Fig. 2c), which instead provided significant quantities of by-products **16** and **17**. Over the course of the project, it was found that Fe(dibm)₃ provided the highest yields when the heteroatom substitution on the donor olefin contained Lewis-basic lone pairs,

¹Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.

*These authors contributed equally to this work.

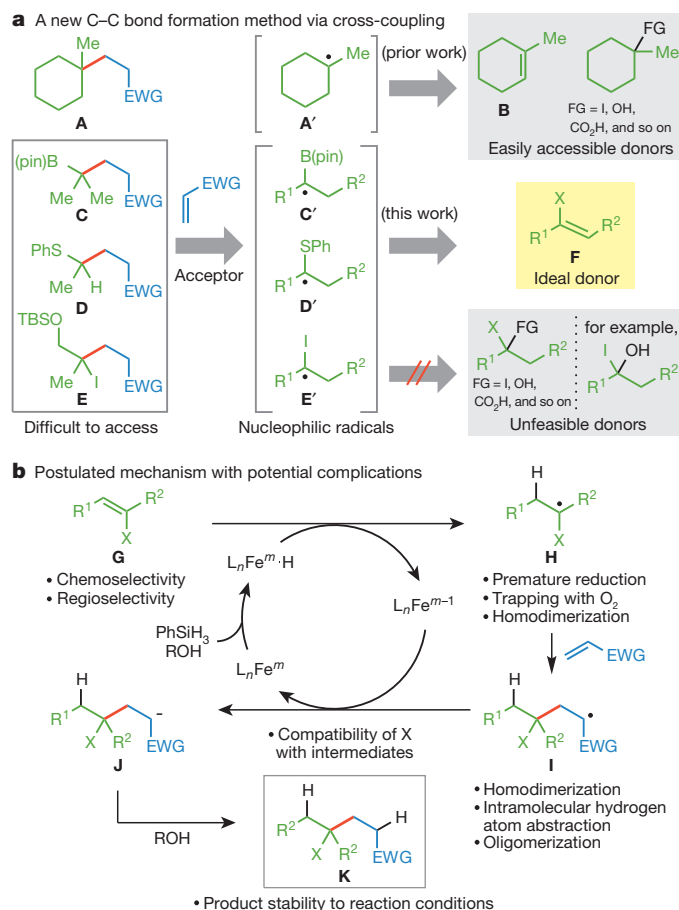


Figure 1 | Functionalized olefin cross-coupling as a strategy for convergent chemical synthesis. **a**, Functionalized olefin cross-coupling would facilitate the exploration of chemical space that has previously been difficult to access (for example, C–E). Such a strategy would use readily available heteroatom-substituted olefins as donors (F) to access nucleophilic radical intermediates (for example, C'–E'), which would couple with electrophilic acceptor olefins. This approach would avoid difficulties that could arise from the use of other radical precursors (greyed box, bottom right). **b**, The functionalized olefin cross-coupling would occur by the Fe hydride-mediated conversion of the donor olefin G to the nucleophilic radical H, which would undergo conjugate addition to the acceptor olefin to form intermediate I. Single-electron reduction to form the stabilized anion J followed by protonation would form the final product K. Examination of the postulated mechanism for the cross-coupling reveals several potential complications (bulleted) that could arise due to either the intermediacy of radicals or the heteroatom (X) present on the donor olefin. EWG, electron-withdrawing group; FG, functional group; (pin), pinacolato; TBS, *tert*-butyldimethylsilyl; X, heteroatom; L, ligand.

whereas Fe(acac)₃ proved superior in the absence of such moieties (see below).

Scope and functional group tolerance

The optimized conditions were then applied to a wider variety of donor and acceptor olefins, initially focusing on enol ethers (Fig. 3a). Using Fe(dibm)₃ (5 mol%), silyl enol ethers could be coupled to cyclic and acyclic enones, an enal and an acrylamide to generate adducts **3** and **20–25** with yields that generally increased with decreasing substitution on the silicon atom (**19** and **22–24**). Remarkably, even a severely congested oestrone derivative could undergo addition to methyl vinyl ketone to generate steroidal adduct **25** with the stereochemistry of the newly formed neopentyl quaternary stereocentre corresponding to that obtained through a conventional organometallic addition of an alkyl group to oestrone²³. Alkyl and aryl vinyl ethers could also be used, although higher yields were generally obtained by using the donor olefin in excess (**26–33**).

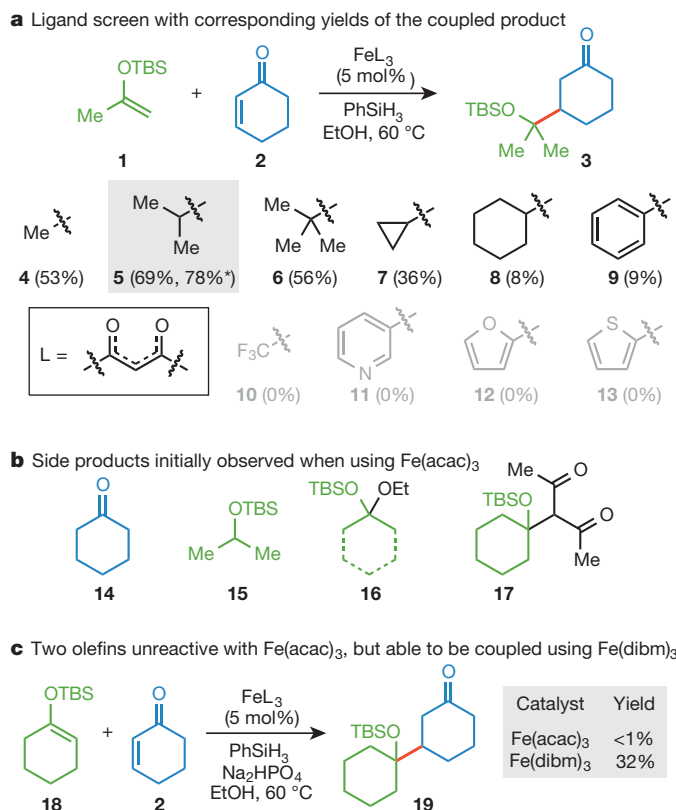


Figure 2 | Functionalized olefin cross-coupling optimization studies. **a**, Top row, reaction studied (ligand L shown bottom left). Altering the ligands on the Fe centre (by using compounds **4–13**) had the greatest influence on the outcome of the reaction, with Fe(dibm)₃ (**5**) giving the highest yields. The addition of 1 equiv. Na₂HPO₄ further increased the yield. (Yields here and in **c** are based on GC/MS analysis using 1,3,5-trimethoxybenzene as an internal standard.) Greyed-out ligands gave 0% yield. **b**, Side products that were observed when Fe(acac)₃ (**4**) was used as the catalyst. The formation of compounds **16** and **17** could be attributed to the Lewis acidity of **4**. The use of **5** as the catalyst reduced the formation of compounds **16** and **17**. **c**, An example where the use of **5** instead of **4** was essential in obtaining the desired functionalized olefin cross-coupling reactivity. TBS, *tert*-butyldimethylsilyl; L, ligand; acac, acetylacetonate; dibm, diisobutyrylmethane; GC/MS, gas chromatography/mass spectrometry.

Endocyclic enol ethers were also tolerated, as shown by the formation of **30–33**.

Additionally, enecarbamates and enamides could undergo cross-coupling under the reaction conditions (Fig. 3b). Adducts **34** and **35** were formed by the coupling of a Cbz (benzyloxycarbonyl)-protected dihydropyrrole with benzyl acrylate and cyclopent-2-enone, respectively, although these couplings necessitated larger amounts of PhSiH₃ than the enol ethers. The amount of PhSiH₃ needed could be decreased by using more electronically activated acceptors, as the formation of **36** and **37** demonstrated. Other cyclic and acyclic enecarbamates could also be employed and added to various acceptor olefins (**38, 39** and **41–46**), although higher loadings (15 mol%) of Fe(dibm)₃ were typically required for useful yields. The formation of **40** also demonstrated that the nitrogen atom present on the donor olefin could be protected as an amide instead of a carbamate. Mono- and 1,1-disubstituted acyclic donor olefins were competent donors (**41–46**), however attempts to control the stereochemistry of the cross-coupling by using α -phenylethylamine as a chiral auxiliary²⁴ provided only modest amounts of diastereoselectivity (**45** and **46**).

Vinyl thioethers proved to be unique donor olefins, with the cross-couplings of those surveyed taking place at ambient temperature to generate adducts **47–56** (Fig. 3c). Although the cross-coupling to form **49**

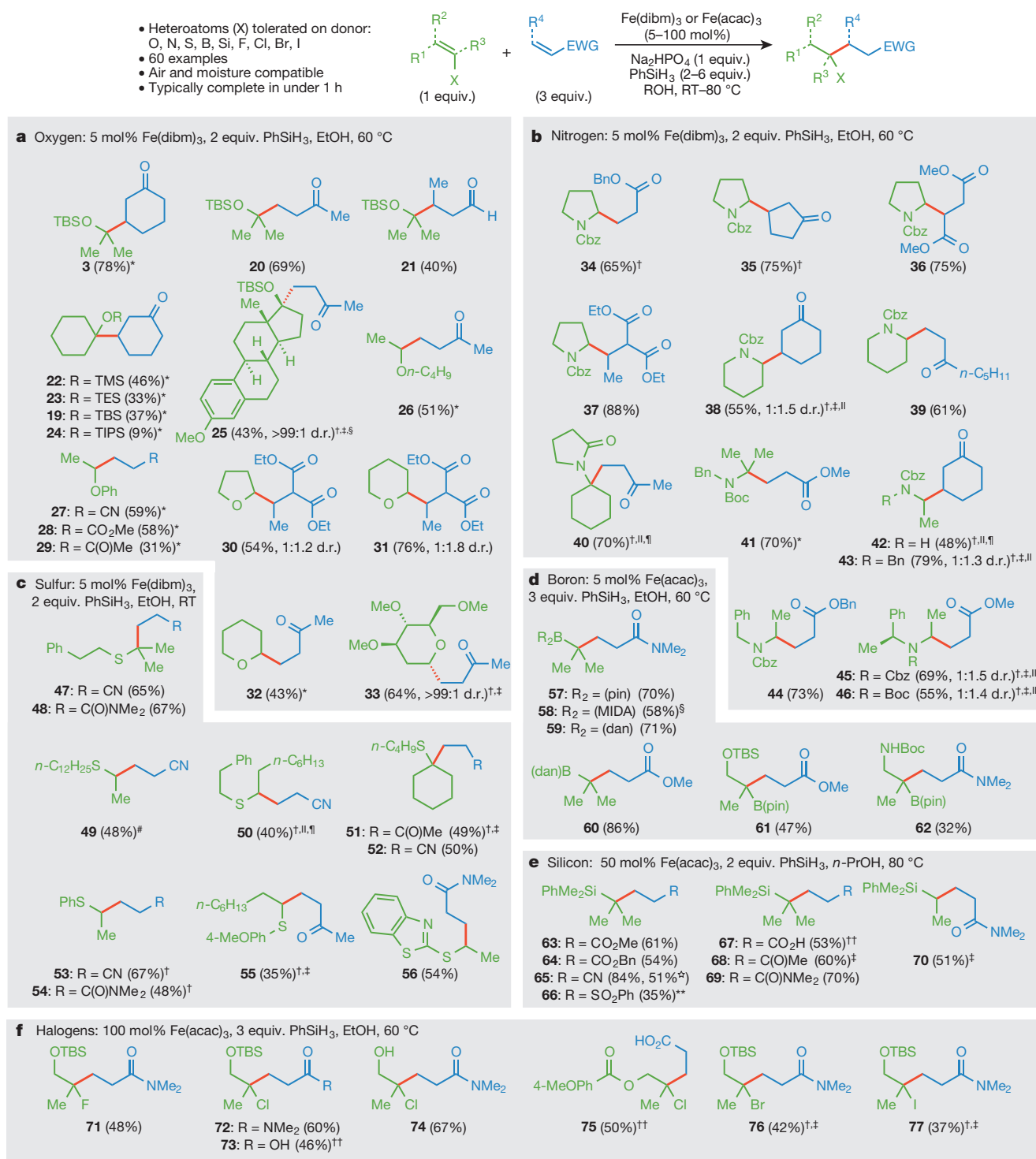


Figure 3 | Adducts synthesized by functionalized olefin cross-coupling. Top, the reaction studied. The donor component is shown in green and the acceptor component is shown in blue. Couplings using donor olefins with heteroatom substitution containing Lewis basic lone pairs (**a**, O; **b**, N; **c**, S) proceeded in higher yields with Fe(dibm)₃ whereas couplings without such moieties (**d**, B; **e**, Si; **f**, halogens) proceeded in higher yields with Fe(acac)₃. *3 equiv. donor and 1 equiv. acceptor used. †6 equiv. PhSiH₃ used. ‡6 equiv.

acceptor used. §THF used as a cosolvent. || 15 mol% [Fe] used. ¶Second portion of [Fe], acceptor and PhSiH₃ added after 1 h. #Heated at 60 °C. ☆Run on gram-scale. **100 mol% [Fe] used. ††Na₂HPO₄ omitted. TMS, trimethylsilyl; TES, triethylsilyl; TIPS, triisopropylsilyl; Bn, benzyl; Cbz, benzyloxycarbonyl; Boc, *tert*-butoxycarbonyl; (MIDA), *N*-methyliminodiacetate; (dan), 1,8-diaminonaphthyl.

proceeded in a higher yield when the reaction was heated at 60 °C, the yields of the other vinyl thioether cross-couplings did not benefit from elevated temperatures. With the exception of **50**, the coupling of the alkenyl thioether donors proceeded with 5 mol% of Fe(dibm)₃; however, increased amounts of PhSiH₃ and acceptor olefin were required

for certain recalcitrant substrates (**50**, **51**, **53** and **54**). Syringe pump addition of the acceptor and PhSiH₃ to the reaction mixture could also improve yields in certain cases (**51** and **55**).

Boron substitution on the donor olefin could also be tolerated, with the use of 5 mol% Fe(acac)₃ providing slightly higher yields than Fe(dibm)₃.

An isopropenyl pinacolato (pin) boronic ester, *N*-methyliminodiacetate (MIDA) boronate^{25,26}, and a 1,8-diaminonaphthyl (dan) boronamide²⁷ could all be coupled to *N,N*-dimethyl acrylamide (**57–59**; Fig. 3d), although the use of THF as a cosolvent was required to solubilize the MIDA boronate. Additionally, methyl acrylate could be used as an acceptor (**60** and **61**), and oxygen- and nitrogen-containing functionalities could be tolerated at allylic positions (**61** and **62**).

Vinyl silanes could also be used as donor olefins, although highest yields were obtained using a substoichiometric amount (50 mol%) of Fe(acac)₃. Additionally, switching the solvent from EtOH to *n*-PrOH and heating the reactions to 80 °C instead of 60 °C resulted in higher yields. With these slight modifications, an isopropenyl and vinyl silane could be coupled to a wide variety of acceptor olefins to form **63–70** (Fig. 3e), although the coupling to obtain the phenyl vinyl sulfone adduct **66** required a stoichiometric amount of Fe(acac)₃. With the omission of Na₂HPO₄, unprotected acrylic acid could be used as an acceptor to provide the coupled product **67** in a transformation difficult to achieve using conventional conjugate addition techniques^{28,29}.

As a final testament to the mildness of this C–C bond forming reaction, alkenyl halides were found to take part in the cross-coupling in reasonable yields using stoichiometric amounts of Fe(acac)₃. Alkenyl fluorides, chlorides, bromides and even iodides could all be used as donors, with the 2-haloallyl alcohol derivatives delivering products **71**, **72**, **76** and **77** (Fig. 3f), where the halogen atom remained intact. Interestingly, acrylic acid could once again be used as an acceptor (**73**, **75**), and the reaction proceeded readily with a free alcohol (**74**), demonstrating the notable chemoselectivity of this method.

To highlight the efficiency of the newly developed coupling reaction, we chose to target glucal derivative **79** (Fig. 4a). This compound has previously been prepared in three steps from readily available **78** in 52% yield, although that route required the use of excess gaseous HCl, toxic and harsh organometallic reagents and cryogenic temperatures³⁰. By contrast, olefin cross-coupling allowed the desired product **79** to be synthesized directly from **78** in a single step over two hours in 68% isolated yield, although it did require the slow addition of a large excess (12 equiv.) of both methyl vinyl ketone and PhSiH₃.

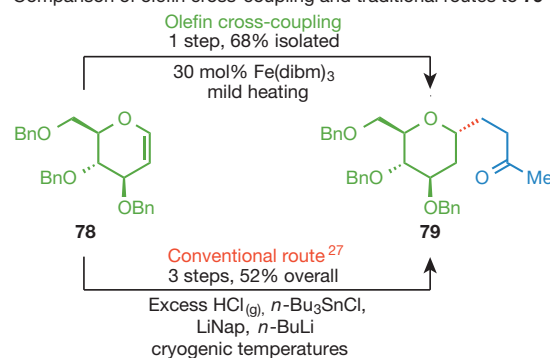
Finally, the resilience of the functionalized olefin cross-coupling to adverse conditions was evaluated by performing the reaction in a variety of unconventional solvents. As indicated by GC/MS, the coupling to form silyl ether **20** proved to be successful in a selection of beer, wine and various spirits (see Supplementary Table 2 and Supplementary Figs 23–30). In addition to showing the ability of the reaction to proceed under aqueous conditions, these results demonstrate the reaction's tolerance of a host of organic compounds³¹ and microorganisms, suggesting possible downstream applications to the area of bioconjugation³².

Discussion and limitations

From a strategic perspective, this methodology grants access to areas of chemical space that, in most cases, were previously inaccessible. Historically, heteroatom-substituted quaternary centres have been synthesized with multiple FG manipulations and rarely, if ever, through a direct C–C disconnection as enabled here. Thus, ~90% of the compounds listed in Fig. 3 are new chemical entities despite their simplicity. In the case of **30**, **31** and **34–37**, where a comparison to contemporary reactivity modes could be made, it was found that the olefin cross-coupling route offers a complementary approach to the recently reported decarboxylative method³³. Furthermore, the olefin cross-coupling reaction set-up was operationally simple, as no precautions were made with regards to moisture or air exclusion, and reactions were typically done within a few minutes to an hour. The reaction is also readily scalable, with the coupling to form **65** being conducted on the gram scale (51% yield).

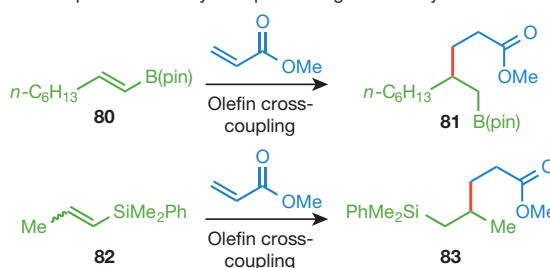
However, no reaction is without limitations. Although nearly all of the substrate classes tested delivered the expected product, the 1,2-disubstituted vinyl boronic ester **80** and vinyl silane **82** exclusively provided adducts **81** and **83**, respectively, where bond formation occurred distal to the heteroatom (Fig. 4b). Additionally, excessive alkyl substitution

a Comparison of olefin cross-coupling and traditional routes to **79**

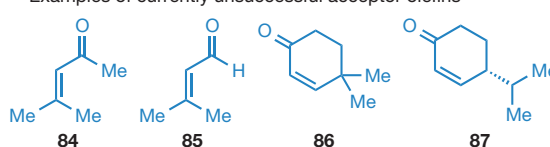


b Limitations of olefin cross-coupling

Examples of currently unexpected regiochemistry



Examples of currently unsuccessful acceptor olefins



c Mechanistic studies

Radical clock experiment



Labelling studies

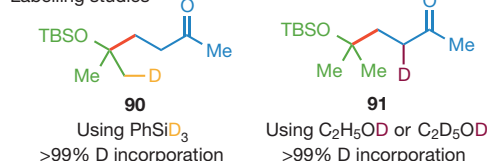


Figure 4 | Additional functionalized olefin cross-coupling studies.

a, Functionalized olefin cross-coupling (top route) offers a direct route to glucal derivative **79** that circumvents the harsh reagents, superstoichiometric organometallic reagents and cryogenic temperatures used in conventional approaches (bottom route). **b**, Top two rows: the use of certain 1,2-disubstituted donor olefins (**80** and **82**) gave adducts where the C–C bond formed distal instead of adjacent to the heteroatom (**81** and **83**). Bottom row: the use of acceptors with excessive aliphatic substitution (**84–87**) gave trace or no product. **c**, The use of vinyl cyclopropane **88** resulted in the isolation of **89**, where the fragmentation of the cyclopropane “radical clock” supports the formation of a radical adjacent to the heteroatom in the donor. Isolation of compounds **90** and **91** from deuterium labelling studies further support the mechanism depicted in Fig. 1b.

on the acceptor olefin was not well tolerated, with trisubstituted acceptors (for example, **84** and **85**) and disubstituted acceptors containing aliphatic β branching (for example, **86** and **87**) generally giving little or no product. Cases where the isolated yield was ~50% and below could be attributed to incomplete conversion, premature reduction or substrate

dimerization. It is finally worth noting that as Fig. 3 demonstrates, the stereochemical outcomes of this reaction are all currently substrate-controlled.

Although a thorough mechanistic investigation has not been pursued, several observations are consistent with the mechanism depicted in Fig. 1b. Subjecting a donor olefin bearing a vinylcyclopropane (**88**, Fig. 4c) to the reaction conditions led to the isolation of adduct **89**, arising from cleavage of the cyclopropane ring. Furthermore, the utilization of PhSiD_3 instead of PhSiH_3 resulted in the isolation of C6 deuterated adduct **90**. These two observations support the notion that a hydrogen atom originating from PhSiH_3 is incorporated into donor olefin **G** (Fig. 1b) through a radical-based process. Boger has previously proposed a similar initiating step in his Fe-mediated oxidation of anhydrovinblastine to vinblastine and originated the idea that Fe-mediated Mukaiyama-type hydrofunctionalizations may not occur via hydrometallation³⁴. In recent work developing a mild thermodynamic olefin reduction applicable to haloalkenes, Shenvi has suggested hydrogen atom transfer (HAT) to be the initial step of these hydrofunctionalizations¹⁸. Taken together, these observations support the initiation of the functionalized olefin cross-coupling by HAT from an Fe hydride³⁵ generated *in situ* to the donor olefin **G** to form radical intermediate **H** (Fig. 1b). The protonation of intermediate **J** to the final coupled product **K** is supported by the isolation of adduct **91** (Fig. 4c) when using either ethanol- d_1 or ethanol- d_6 as the solvent. Submitting undeuterated analogue **20** (Fig. 3a) to the reaction conditions using deuterated ethanol did not lead to any deuterium incorporation, demonstrating that the deuterium incorporation observed in the labelling studies occurred during the course of the reaction.

Conclusion

In summary, a new method for forming unique C–C bonds in a rapid, scalable and practical fashion has been described using an inexpensive iron catalyst and a simple reaction set-up. From a retrosynthetic perspective, this method requires the rethinking of the classic roles of some common building blocks in organic synthesis. For example, enol ethers and enamides need not be viewed as reacting as nucleophiles solely at their β position^{36,37}. Vinyl boronates, normally used to fashion new $\text{C}(\text{sp}^2)$ centres³⁸, can now be viewed as potential progenitors to tertiary boronates for a variety of Ni- and Pd-based $\text{C}(\text{sp}^3)$ couplings³⁹. Vinyl thioethers, rarely employed in molecule construction⁴⁰, can now be viewed

in a different light. Vinyl silanes have been employed in cyclizations⁴¹ and $\text{C}(\text{sp}^2)$ cross-coupling chemistry⁴² but never as precursors to silyl-substituted quaternary centres. In the case of vinyl halides, the halide (F, Cl, Br and even I) no longer needs to be viewed as a disposable functionality for conventional transition-metal-mediated cross-coupling⁴³, but rather as a spectator FG that can be incorporated into a final product. Functionalized olefin cross-coupling ultimately represents a method of reversing the native reactivity⁴⁴ of heteroatom-substituted olefins (Fig. 5), thus permitting the facile exploration of underdeveloped chemical space and serving as an alternative to other powerful retrosynthetic C–C bond disconnections^{45–47}. Although achieving ligand control of stereo- and regiochemical outcomes and a deeper understanding of the mechanism are prominent future goals, potential applications of this method, even in its current form, to numerous areas of chemical science can be envisioned.

Received 13 September; accepted 20 October 2014.

- Corey, E. J. & Cheng, X.-M. *The Logic of Chemical Synthesis* (Wiley, 1995).
- Isayama, S. & Mukaiyama, T. A new method for the preparation of alcohols from olefins with molecular oxygen and phenylsilane by the use of bis(acetylacetonato)cobalt(II). *Chem. Lett.* **18**, 1071–1074 (1989).
- Kato, K. & Mukaiyama, T. Iron(III) complex catalyzed nitrosation of terminal and 1,2-disubstituted olefins with butyl nitrite and phenylsilane. *Chem. Lett.* **21**, 1137–1140 (1992).
- Waser, J., Gaspar, B., Nambu, H. & Carreira, E. M. Hydrazines and azides via the metal-catalyzed hydrohydrazination and hydroazidation of olefins. *J. Am. Chem. Soc.* **128**, 11693–11712 (2006).
- Leggans, E. K., Barker, T. J., Duncan, K. K. & Boger, D. L. Iron(III)/ NaBH_4 -mediated additions to unactivated alkenes: synthesis of novel 20'-vinblastine analogues. *Org. Lett.* **14**, 1428–1431 (2012).
- Shigehisa, H., Aoki, T., Yamaguchi, S., Shimizu, N. & Hiroya, K. Hydroalkoxylation of unactivated olefins with carbon radicals and carbocation species as key intermediates. *J. Am. Chem. Soc.* **135**, 10306–10309 (2013).
- Shigehisa, H., Nishi, E., Fujisawa, M. & Hiroya, K. Cobalt-catalyzed hydrofluorination of unactivated olefins: a radical approach of fluorine transfer. *Org. Lett.* **15**, 5158–5161 (2013).
- Girijavallabhan, V., Alvarez, C. & Njoroge, F. G. Regioselective cobalt-catalyzed addition of sulfides to unactivated alkenes. *J. Org. Chem.* **76**, 6442–6446 (2011).
- Taniguchi, T., Goto, N., Nishibata, A. & Ishibashi, H. Iron-catalyzed redox radical cyclizations of 1,6-dienes and enynes. *Org. Lett.* **12**, 112–115 (2010).
- Wang, L. C. *et al.* Diastereoselective cycloreductions and cycloadditions catalyzed by $\text{Co}(\text{dpm})_2$ -silane (dpm) 2,2,6,6-tetramethylheptane-3,5-dionate: mechanism and partitioning of hydrometallative versus anion radical pathways. *J. Am. Chem. Soc.* **124**, 9448–9453 (2002).
- Streuff, J. The electron-way: metal-catalyzed reductive umpolung reactions of saturated and α,β -unsaturated carbonyl derivatives. *Synthesis* **45**, 281–307 (2013).
- Zbieg, J. R., Yamaguchi, E., McInturff, E. L. & Krische, M. J. Enantioselective C–H crotylation of primary alcohols via hydrohydroxyalkylation of butadiene. *Science* **336**, 324–327 (2012).
- Lo, J. C., Yabe, Y. & Baran, P. S. A practical and catalytic reductive olefin coupling. *J. Am. Chem. Soc.* **136**, 1304–1307 (2014).
- Srikanth, G. S. C. & Castle, S. L. Advances in radical conjugate additions. *Tetrahedron* **61**, 10377–10441 (2005).
- Lackner, G. L., Quasdorf, K. W. & Overman, L. E. Direct construction of quaternary carbons from tertiary alcohols via photoredox-catalyzed fragmentation of *tert*-alkyl *N*-phthalimidoyl oxalates. *J. Am. Chem. Soc.* **135**, 15342–15345 (2013).
- Barton, D. H. R. & Crich, D. Formation of quaternary carbon centres from tertiary alcohols by free radical methods. *Tetrahedron Lett.* **26**, 757–760 (1985).
- Barton, D. H. R., Crich, D. & Kretschmar, G. Formation of carbon-carbon bonds with radicals derived from the esters of thiohydroxamic acids. *Tetrahedron Lett.* **25**, 1055–1058 (1984).
- Iwasaki, K., Wan, K. W., Oppedisano, A., Crossley, S. W. M. & Shenvi, R. A. Simple, chemoselective hydrogenation with thermodynamic stereocontrol. *J. Am. Chem. Soc.* **136**, 1300–1303 (2014).
- King, S. M., Ma, X. & Herzog, S. B. A method for the selective hydrogenation of alkenyl halides to alkyl halides. *J. Am. Chem. Soc.* **136**, 6884–6887 (2014).
- Magnus, P., Waring, M. J. & Scott, D. A. Conjugate reduction of α,β -unsaturated ketones using an Mn^{III} catalyst, phenylsilane and isopropyl alcohol. *Tetrahedron Lett.* **41**, 9731–9733 (2000).
- Zotto, C. D. *et al.* FeCl_3 -catalyzed addition of nitrogen and 1,3-dicarbonyl nucleophiles to olefins. *J. Organomet. Chem.* **696**, 296–304 (2011).
- Shigematsu, T., Matsui, M. & Utsunomiya, K. Gas chromatography of diisobutylmethane metal chelates. *Bull. Inst. Chem. Res. Kyoto Univ.* **46**, 256–261 (1968).
- Djerassi, C., Miramontes, L., Rosenkranz, G. & Sondheimer, F. Steroids. LIV. Synthesis of 19-nor-17 α -ethynyltestosterone and 19-nor-17 α -methyltestosterone. *J. Am. Chem. Soc.* **76**, 4092–4094 (1954).
- Juaristi, E., León-Romo, J. L., Reyes, A. & Escalante, J. Recent applications of α -phenylethylamine (α -PEA) in the preparation of enantiopure compounds. Part 3: α -PEA as chiral auxiliary. Part 4: α -PEA as chiral reagent in the

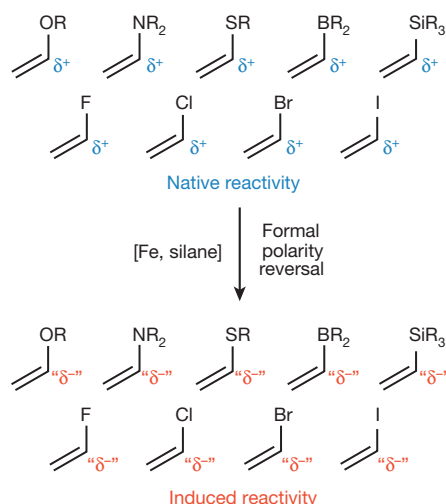


Figure 5 | Functionalized olefin cross-coupling reverses conventional reactivity expectations. The substrates employed as donors in this study typically are electrophilic (δ^+) at the position adjacent to the heteroatom. Functionalized olefin cross-coupling reverses this native reactivity by generating radical intermediates through the use of an Fe catalyst and a silane. These radicals induce nucleophilic properties (δ^-) at those formerly electrophilic positions, resulting in a reversal in typical reactivity.

- stereodifferentiation of prochiral substrates. *Tetrahedron Asymmetry* **10**, 2441–2495 (1999).
25. Mancilla, T. & Contreras, R. New bicyclic organylboronic esters derived from iminodiacetic acids. *J. Organomet. Chem.* **307**, 1–6 (1986).
 26. Gillis, E. P. & Burke, M. D. A simple and modular strategy for small molecule synthesis: iterative Suzuki–Miyaura coupling of B-protected haloboronic acid building blocks. *J. Am. Chem. Soc.* **129**, 6716–6717 (2007).
 27. Noguchi, H., Hojo, K. & Sugimoto, M. Boron-masking strategy for the synthesis of oligoarenes via iterative Suzuki–Miyaura coupling. *J. Am. Chem. Soc.* **129**, 758–759 (2007).
 28. Yamamoto, Y. & Maruyama, K. $\text{RCu}\cdot\text{BF}_3$. 3. Conjugate addition to previously unreactive substituted enoate esters and enoic acids. *J. Am. Chem. Soc.* **100**, 3240–3241 (1978).
 29. Aurell, M. J., Domingo, L. R., Mestres, R., Muñoz, E. & Zaragová, R. J. Conjugate addition of organolithium reagents to α,β -unsaturated carboxylic acids. *Tetrahedron* **55**, 815–830 (1999).
 30. Hutchinson, D. K. & Fuchs, P. L. Amelioration of the conjugate addition chemistry of α -alkoxycopper reagents: application to the stereospecific synthesis of C-glycosides. *J. Am. Chem. Soc.* **109**, 4930–4939 (1987).
 31. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Vol. 44, *Alcohol Drinking* 71–99 (World Health Organization, 1988).
 32. Kolb, H. C. & Sharpless, K. B. The growing impact of click chemistry on drug discovery. *Drug Discov. Today* **8**, 1128–1137 (2003).
 33. Chu, L., Ohta, C., Zuo, Z. & MacMillan, D. W. C. Carboxylic acids as a traceless activation group for conjugate additions: a three-step synthesis of (\pm)-pregabalin. *J. Am. Chem. Soc.* **136**, 10886–10889 (2014).
 34. Ishikawa, H. *et al.* Total synthesis of vinblastine, vincristine, related natural products, and key structural analogs. *J. Am. Chem. Soc.* **131**, 4904–4916 (2009).
 35. Bullock, R. M. & Samsel, E. G. Hydrogen atom transfer reactions of transition-metal hydrides. Kinetics and mechanism of the hydrogenation of α -cyclopropylstyrene by metal carbonyl hydrides. *J. Am. Chem. Soc.* **112**, 6886–6898 (1990).
 36. Matsuo, J. & Murakami, M. The Mukaiyama aldol reaction: 40 years of continuous development. *Angew. Chem. Int. Edn* **52**, 9109–9118 (2013).
 37. Stork, G., Brizzolara, A., Landesman, H., Szmuszkovicz, J. & Terrell, R. The enamine alkylation and acylation of carbonyl compounds. *J. Am. Chem. Soc.* **85**, 207–222 (1963).
 38. Miyaura, N. & Suzuki, A. Palladium-catalyzed cross-coupling reactions of organoboron compounds. *Chem. Rev.* **95**, 2457–2483 (1995).
 39. Jana, R., Pathak, T. P. & Sigman, M. S. Advances in transition metal (Pd, Ni, Fe)-catalyzed cross-coupling reactions using alkyl-organometallics as reaction partners. *Chem. Rev.* **111**, 1417–1492 (2011).
 40. Dubbaka, S. R. & Vogel, P. Organosulfur compounds: electrophilic reagents in transition-metal-catalyzed carbon–carbon bond-forming reactions. *Angew. Chem. Int. Edn* **44**, 7674–7684 (2005).
 41. Blumenkopf, T. A. & Overman, L. E. Vinylsilane- and alkynylsilane-terminated cyclization reactions. *Chem. Rev.* **86**, 857–873 (1986).
 42. Nakao, Y. & Hiyama, T. Silicon-based cross-coupling reaction: an environmentally benign version. *Chem. Soc. Rev.* **40**, 4893–4901 (2011).
 43. Diederich, F. & Stang, P. J. (eds) *Metal-catalyzed Cross-coupling Reactions* (Wiley-VCH, 1998).
 44. Seebach, D. Methods of reactivity umpolung. *Angew. Chem. Int. Edn Engl.* **18**, 239–258 (1979).
 45. Gao, X., Soo, S. K. & Krische, M. J. Total synthesis of 6-deoxyerythronolide B via C–C bond-forming transfer hydrogenation. *J. Am. Chem. Soc.* **135**, 4223–4226 (2013).
 46. Werner, E. W., Mei, T.-S., Burckle, A. J. & Sigman, M. S. Enantioselective Heck arylations of acyclic alkenyl alcohols using a redox-relay strategy. *Science* **338**, 1455–1458 (2012).
 47. Meek, S. J., O'Brien, R. V., Llaveria, J., Schrock, R. J. & Hoveyda, A. H. Catalytic Z-selective olefin cross-metathesis for natural product synthesis. *Nature* **471**, 461–466 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements Financial support for this work was provided by NIH/NIGMS (GM-097444). The National Science Foundation supported a predoctoral fellowship for J.C.L.; the Shanghai Institute of Organic Chemistry, Zhejiang Medicine Co. and Pharmaron supported a postdoctoral fellowship for J.G.; and the Japan Society for the Promotion of Science supported a postdoctoral fellowship for Y.Y. We are grateful to D.-H. Huang and L. Pasternack (TSRI) for assistance with NMR spectroscopy, and A. L. Rheingold and C. E. Moore (UCSD) for X-ray crystallographic analysis. We thank R. A. Shenvi (TSRI) and Y. Ji (TSRI) for discussions.

Author Contributions J.C.L. and P.S.B. conceived the work; J.C.L. conducted initial feasibility studies; J.C.L., J.G., Y.Y., C.-M.P. and P.S.B. designed the experiments and analysed the data; J.C.L., J.G., Y.Y. and C.-M.P. performed the experiments; and J.C.L. and P.S.B. wrote the manuscript.

Author Information Crystallographic data for the structure of $\text{Fe}(\text{dibm})_3$ (**5**) is available free of charge from the Cambridge Crystallographic Data Centre under deposition number CCDC 1022625. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.S.B. (pbaran@scripps.edu).

An AUTS2–Polycomb complex activates gene expression in the CNS

Zhonghua Gao¹, Pedro Lee¹, James M. Stafford¹, Melanie von Schimmelmann², Anne Schaefer² & Danny Reinberg¹

Naturally occurring variations of Polycomb repressive complex 1 (PRC1) comprise a core assembly of Polycomb group proteins and additional factors that include, surprisingly, autism susceptibility candidate 2 (AUTS2). Although AUTS2 is often disrupted in patients with neuronal disorders, the mechanism underlying the pathogenesis is unclear. We investigated the role of AUTS2 as part of a previously identified PRC1 complex (PRC1–AUTS2), and in the context of neurodevelopment. In contrast to the canonical role of PRC1 in gene repression, PRC1–AUTS2 activates transcription. Biochemical studies demonstrate that the CK2 component of PRC1–AUTS2 neutralizes PRC1 repressive activity, whereas AUTS2-mediated recruitment of P300 leads to gene activation. Chromatin immunoprecipitation followed by sequencing (ChIP-seq) demonstrated that AUTS2 regulates neuronal gene expression through promoter association. Conditional targeting of *Auts2* in the mouse central nervous system (CNS) leads to various developmental defects. These findings reveal a natural means of subverting PRC1 activity, linking key epigenetic modulators with neuronal functions and diseases.

Polycomb group (PcG) proteins maintain repressive forms of chromatin and therefore appropriate patterns of gene repression through epigenetic mechanisms. As such, PcG proteins have key roles in normal developmental progression, stem cell biology and cancer^{1–8}. The two major groups of PcG protein complexes exhibit distinct enzymatic activities: Polycomb repressive complex 2 (PRC2) catalyses di- and tri-methylation of histone H3 at lysine 27 (H3K27me2/3)^{9–12}, and Polycomb repressive complex 1 (PRC1) catalyses monoubiquitination of histone H2A at lysine 119 (H2AK119ub1)^{13,14} and/or compacts chromatin¹⁵. There are at least six distinct groups of mammalian PRC1 complexes, PRC1.1–1.6, each comprising one of six Polycomb group RING fingers (PCGFs)¹⁶, and the E3 ligase RING1A/B. Further diversification arises from the mutually exclusive association of RING1A/B with either RYBP or YAF2, or one of the CBX proteins^{16–18}, which bind H3K27me3 through their chromodomains. Unlike their CBX-containing counterparts, RYBP-containing PRC1 complexes adopt a PRC2/H3K27me3-independent mechanism for targeting chromatin¹⁷.

Our previous studies revealed that PCGF3 and PCGF5 form novel PRC1 complexes which contain AUTS2 (ref. 16). *AUTS2* maps to chromosome 7q11.2, encodes a nuclear protein¹⁹, and is frequently reported as being disrupted in individuals suffering neurological disorders, including autism spectrum disorders (ASD)^{20,21}. Although recent studies implicate *auts2* in regulating head size, neurodevelopment and enhancer function in zebrafish^{22,23}, the function of the AUTS2 protein has not been established nor has its role in regulating neuronal functions whose deregulation may result in pathogenesis.

The physical link between PRC1, a key epigenetic regulator, and AUTS2, a risk factor for ASD and other neurological disorders, prompted us to investigate the functional role of the AUTS2-containing PRC1 complex (PRC1–AUTS2). Here we report that PRC1–AUTS2 exhibits an unexpected role in transcriptional activation, in contrast to the repressive role of canonical PRC1. Furthermore, this conversion is mediated by AUTS2. Specific deletion of the *Auts2* locus in mouse neuronal progenitor cells revealed a profound neurodevelopmental phenotype, in accordance with *AUTS2* disruptions in humans.

An AUTS2-containing PRC1 complex

We pursued the unexpected association between PRC1 and AUTS2 (ref. 16) using tandem affinity purification (TAP), followed by mass spectrometry (MS) analysis with AUTS2 fused to sequential N-terminal Flag and HA tags (NFH). As previously reported¹⁶, NFH–AUTS2 was associated with PCGF3, and with components of PRC1.5, including PCGF5, RING1A/B, RYBP and its homologue YAF2, and casein kinase 2 (CK2) (Fig. 1a). We focused on the AUTS2-containing PRC1.5 complex that we designated PRC1.5–AUTS2. Interestingly, several polypeptides that are not PRC1 components, including the co-activator P300, were also associated with AUTS2 (Fig. 1a). Immunoprecipitation (IP) experiments performed with nuclear extract (NE) of 293 T-REx cells expressing a doxycycline-inducible NFH–AUTS2 and antibody against HA confirmed AUTS2 association with RING1B and PCGF5 (Fig. 1b). Other PRC1 components not associated with PRC1.5, such as CBX2, PCGF4 (also known as BMI1), and PCGF1, comprising PRC1.2/4, PRC1.4, and PRC1.1, respectively, did not co-immunoprecipitate with AUTS2 (Fig. 1b). *Auts2* expression at the mRNA level was previously documented in mouse brain via *in situ* hybridization¹⁹. Indeed, RING1B, but not CBX2, interacts with AUTS2 in co-immunoprecipitation experiments performed using nuclear extract of embryonic day 15 (E15) mouse brain and AUTS2 antibody (Fig. 1c), suggesting that PRC1.5–AUTS2 forms within the CNS.

AUTS2, PCGF5, RING1B, CK2B, and RYBP appear to form a stable complex as evidenced by glycerol gradient analysis of AUTS2-containing complexes (Fractions 9–11, Fig. 1d). Although PCGF5 bound both RING1B and AUTS2 (Fig. 1e), RING1B interacted with AUTS2 only in the presence of PCGF5 as evidenced by immunoprecipitations performed *in vitro* using insect-cell-expressed proteins (Fig. 1f). PCGF5 is probably required to bridge RING1B and AUTS2 in complex formation. A similar *in vitro* immunoprecipitation experiment demonstrated that AUTS2 directly interacted with CK2 (Fig. 1g). Thus, PRC1.5–AUTS2 contains at least five components: RYBP/YAF2, RING1A/B, PCGF5, AUTS2, and CK2 (Fig. 1h). RING1A/B associates with PCGF5 as it does with other PCGFs¹⁶, and recruits RYBP/YAF2. AUTS2 is incorporated

¹Howard Hughes Medical Institute, New York University Langone School of Medicine, Department of Biochemistry and Molecular Pharmacology, New York, New York 10016, USA. ²Friedman Brain Institute, Department of Neuroscience, Mount Sinai School of Medicine, New York, New York 10029, USA.

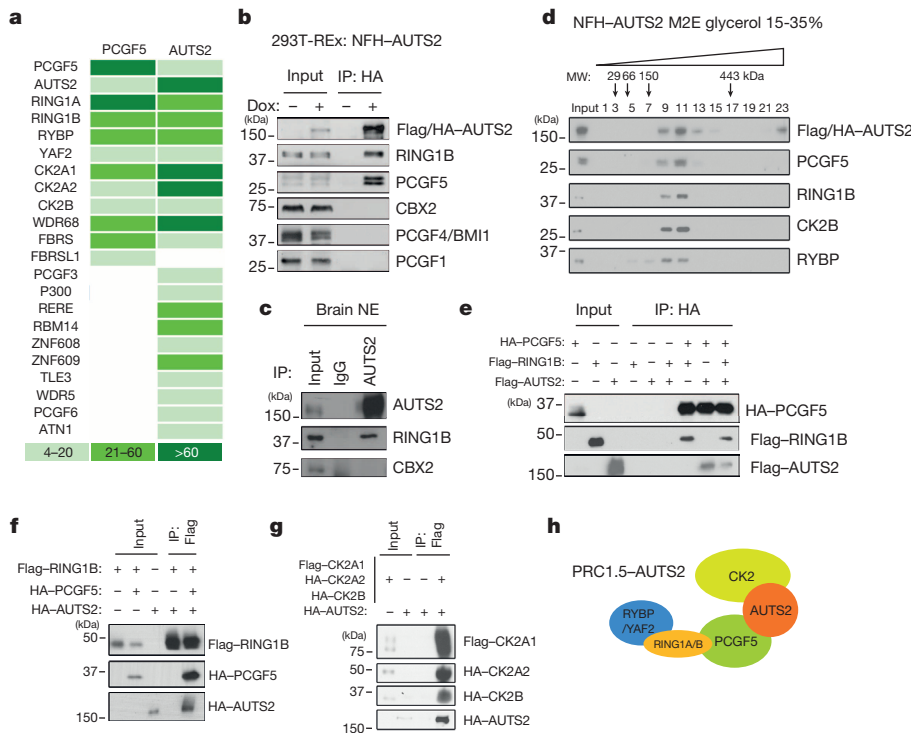


Figure 1 | Characterization of the PRC1.5-AUTS2 complex. **a**, Heat map of PCGF5- and AUTS2-associated polypeptides. Sequential Flag and HA affinity purification (TAP) followed by liquid chromatography–mass spectrometry (LC-MS) (Methods) was performed in inducible 293T-REx cells expressing Flag-HA-tagged bait proteins (PCGF5 TAP from an earlier report)¹⁶ indicated at the top. The spectrum count of each protein is colour-coded and displayed as a heat map. **b**, Immunoprecipitation (IP) from nuclear extract (NE) of NFH-AUTS2 cells in the absence or presence of doxycycline at 100 $\mu\text{g ml}^{-1}$, using HA beads. NFH, N-terminal Flag and HA tags. Bound proteins were resolved on SDS-PAGE and detected by western blotting for the antigens indicated. A total of 5% of input was loaded in all cases, unless indicated otherwise. **c**, Immunoprecipitation from nuclear extract of mouse brain using AUTS2 antibody, followed by western blotting for the antigens indicated. **d**, Glycerol gradient (15–35%) analysis of Flag-purified NFH-AUTS2 complexes. Every other fraction was resolved on SDS-PAGE followed by immunoblotting for the antigens indicated. M2E, M2 beads eluate. **e–g**, Immunoprecipitation with lysates from Sf9 cells expressing the proteins indicated, using HA (**e**) or Flag beads (**f, g**), followed by western blotting using Flag or HA antibodies. **h**, Schematic organization of PRC1.5-AUTS2. See main text for details.

through its interaction with PCGF5. Moreover, AUTS2 recruits CK2 through direct interaction.

PRC1.5-AUTS2 activates transcription

Given that PRC1 functions as a transcriptional repressor through epigenetic mechanisms^{1–3,7}, we examined the effects of PRC1.5-AUTS2 on transcription and chromatin composition. We generated stable 293 T-REx cells containing an integrated luciferase reporter with five consecutive GAL4 DNA binding sites (UAS) (Fig. 2a), and one of the following doxycycline-inducible candidates: GAL4-AUTS2, GAL4-RING1B, GAL4-PCGF4, GAL4-PCGF5 or GAL4 alone. Doxycycline treatment reduced luciferase activity in GAL4-PCGF4 cells, consistent with its role in transcriptional repression (Fig. 2b). Surprisingly, doxycycline treatment increased luciferase activity in GAL4-PCGF5 or GAL4-AUTS2 cells (Fig. 2b; Extended Data Fig. 1a). This result was not due to a post-transcriptional event, as replacing the GAL4 DNA binding sequence with a Flag-HA tag (Flag-HA-AUTS2) resulted in the loss of AUTS2-associated transcriptional activation (Fig. 2b). Interestingly, GAL4-RING1B gave rise to considerably weaker repression, compared to GAL4-PCGF4

(Fig. 2b). Given that RING1A/B comprises all mammalian PRC1 complexes, the net outcome of GAL4-RING1B on transcription probably reflected the sum of all its associated complexes: repressive ones comprising PCGF4, and active ones comprising PCGF5 and AUTS2.

We next probed the luciferase reporter in GAL4-AUTS2 cells for the presence of PRC1.5-AUTS2 components, as well as indicators of chromatin structure, using ChIP followed by quantitative real-time polymerase chain reaction (qPCR). Upon doxycycline induction, GAL4-AUTS2 was recruited to the promoter region of the integrated luciferase locus (Fig. 2c). As expected, RING1B and CK2B were also recruited (Fig. 2c). Consistent with the locus being transcriptionally active (Fig. 2b), Pol II was recruited, accompanied by an increase in trimethylation of histone H3 at lysine 4 (H3K4me3) which correlates with active transcription, and a reduction in H3K27me3 which correlates with repression, without appreciable change in total histone H3 (Fig. 2c). Similar results were obtained using the GAL4-PCGF5 stable line, showing enrichment of Pol II and acetylation of histone H4 at lysine 16 (H4K16ac), an active histone modification, and reduction in H3K27me3 (Fig. 2d). In addition to AUTS2, GAL4-PCGF5 recruited RING1B and CK2B (Fig. 2d).

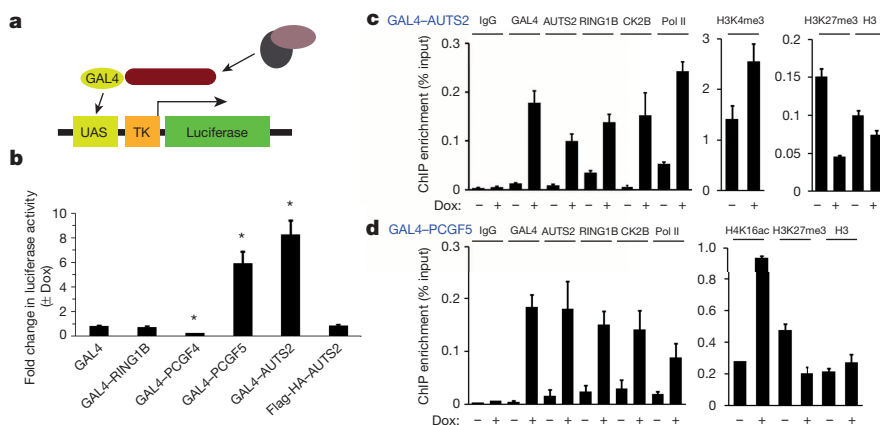


Figure 2 | Effect of AUTS2 on chromatin architecture and transcription. **a**, Schematic of the luciferase reporter system. **b**, Fold change in luciferase activity in cells expressing GAL4, GAL4-RING1B, GAL4-PCGF4, GAL4-PCGF5, GAL4-AUTS2, and Flag-HA-AUTS2, after 24 h doxycycline (100 $\mu\text{g ml}^{-1}$) induction. In this and subsequent luciferase reporter assays, each value is the mean of three independent measurements with error bars representing the standard error. $*P < 0.05$ by two-sided *t*-test, compared with the GAL4 control. **c, d**, ChIP using the antibodies indicated and GAL4-AUTS2 and GAL4-PCGF5 cells, respectively, before and after doxycycline induction. ChIP enrichment is expressed as percentage of input. All centre values represent the average of three technical replicates with error bars indicating standard deviation in this and all other ChIP experiments in this paper.

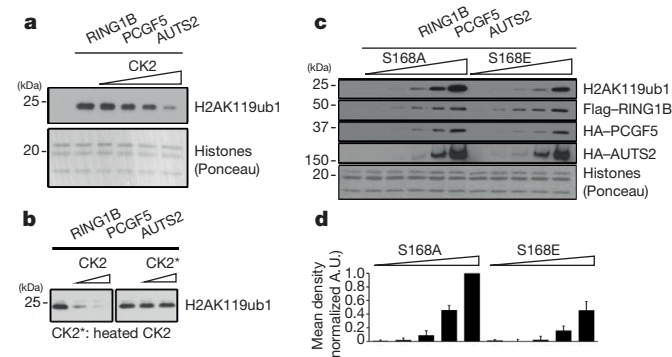


Figure 3 | CK2 inhibits H2A monoubiquitination activity of PRC1.5-AUTS2. **a**, *In vitro* nucleosomal H2A monoubiquitination assay using RING1B-PCGF5-AUTS2 expressed in and purified from Sf9 cells (Methods), as a function of the presence of recombinant CK2. Reaction products were resolved on SDS-PAGE, followed by immunoblotting for H2AK119ub1. Ponceau staining of histones is shown (bottom). **b**, Effect of heat-treatment of CK2 at 95 °C for 15 min before addition to the H2A monoubiquitination assay performed as in **a**. **c**, H2A monoubiquitination assay (Methods) using increasing amounts of RING1B-PCGF5-AUTS2 containing either RING1B(S168A) (S168 to alanine), or RING1B(S168E) (S168 to glutamic acid), purified from Sf9 cells. **d**, Densitometry of H2A monoubiquitination based on three independent experiments as in **c**. Quantification was done by using ImageJ software. Error bars represent standard deviation.

To determine whether reporter gene activation required other components of the PRC1.5-AUTS2 complex, we silenced RING1B or PCGF5 through short interfering RNAs (siRNAs) in GAL4-AUTS2 cells. Indeed, relative to control siRNA treatment, luciferase activity driven by GAL4-AUTS2 was reduced (Extended Data Fig. 1b, c). Similarly, AUTS2 knock-down in GAL4-PCGF5 cells led to decreased luciferase activity (Extended Data Fig. 1d, e). Thus, PRC1.5-AUTS2 creates a chromatin environment favourable to transcription dependent upon the integrity of the complex.

Mechanism of transcriptional activation

PRC1 can monoubiquitinate H2AK119 through its RING1B component. Addition of the PRC1.5-AUTS2 component, CK2, to the recombinant ternary complex RING1B-PCGF5-AUTS2 (Extended Data Fig. 2a), compromised its ability to monoubiquitinate H2AK119 (Fig. 3a). To

rule out the possibility that CK2 may diminish the requisite ATP levels, E1 was pre-charged with ubiquitin and then quenched by EDTA (Extended Data Fig. 2b). Inhibition of H2A monoubiquitination was still evident upon CK2 addition (Extended Data Fig. 2c). Heat inactivation of CK2 obliterated its inhibitory effect (Fig. 3b), indicating that its kinase activity may be required. Indeed, a kinase assay performed with CK2 and components used in the H2A monoubiquitination assay in the presence of [γ - 32 P]ATP demonstrated that CK2 phosphorylated both RING1B and UbcH5c (Extended Data Fig. 2d). An additional radiolabelled species was detectable, dependent upon the presence of nucleosomes, suggesting that CK2 may target one of the core histones under these conditions. Mass spectrometry analysis revealed that CK2 phosphorylated serines 41 (S41) and 168 (S168) of RING1B (Supplementary Table 1). RING1B substitution mutants were generated, containing either aspartic acid (RING1B-SD) or glutamic acid (RING1B-SE) to mimic phosphorylation at S41 and S168, respectively, or alanine (RING1B-SA) to mimic non-phosphorylated forms. Compared to the appropriate alanine substitution, RING1B(S168E) was substantially weaker in H2A monoubiquitination activity, (Fig. 3c, d), whereas RING1B(S41D) showed no obvious difference (Extended Data Fig. 2e). Thus, CK2 inhibits PRC1.5-AUTS2 monoubiquitination of H2AK119 through phosphorylation of RING1B at S168.

Although the presence of CK2 in PRC1.5-AUTS2 suppresses its H2A monoubiquitination activity, robust transcriptional activation probably entails an additional mechanism(s). Of note, the co-activator P300 that facilitates active transcription can associate with AUTS2 in a PRC1-independent manner (Fig. 1a). Co-immunoprecipitation experiments in 293T cells confirmed that AUTS2 and P300 interact and *in vitro* pull-down experiments using recombinant AUTS2 and P300 demonstrated their direct interaction (Extended Data Fig. 3). Importantly, P300 was recruited to the promoter of the luciferase reporter upon GAL4-AUTS2 induction, as evidenced by ChIP results (Fig. 4a). Moreover, siRNA-mediated silencing of P300 (Fig. 4b, insert) led to a dramatic loss in GAL4-AUTS2-mediated activation of the luciferase reporter (Fig. 4b; Extended Data Fig. 4a). A similar loss of activation was obtained upon treating the cells with the C646 inhibitor that specifically blocks P300 acetyltransferase activity (Fig. 4c; Extended Data Fig. 4b). On the basis of results using GAL4 fusions with truncated versions of AUTS2 *in vivo*, AUTS2 amino acids 404 to 913 (GAL4-AUTS2M) was sufficient to mediate transcriptional activation (Fig. 4d). Interestingly, this region overlaps a C-terminal portion of AUTS2 that correlates with severe human phenotypes²². ChIP analysis revealed that truncated versions of

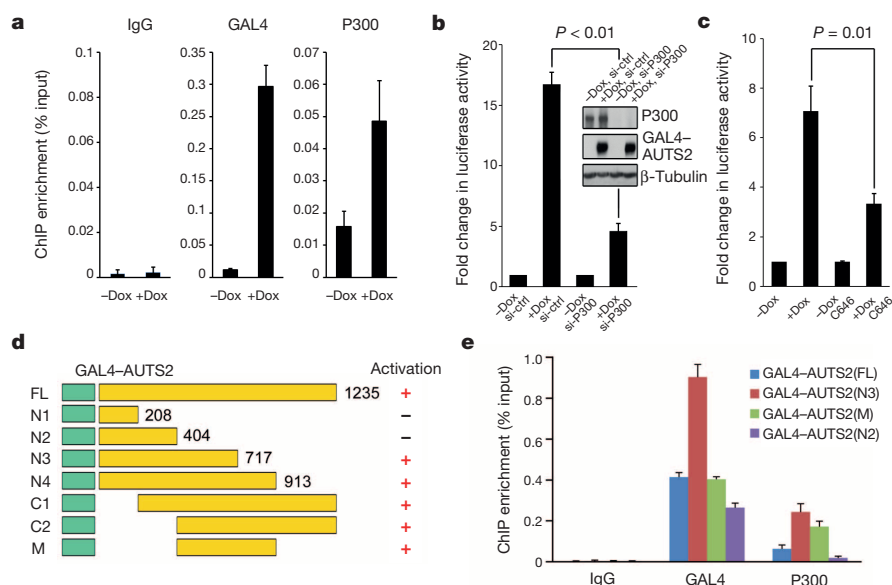


Figure 4 | AUTS2 recruits P300 for gene activation. **a**, ChIP analysis of P300 in GAL4-AUTS2 cells, before and after doxycycline induction, as indicated. **b**, Fold change in luciferase activity in GAL4-AUTS2 cells with control (si-ctrl) or P300 siRNA (si-P300) treatment for 48 h, before 24 h doxycycline induction. Insert shows immunoblotting of samples used for luciferase reporter assays using the antibodies indicated. The *P* value derived from a two-sided *t*-test is indicated here and in **c**. **c**, Fold change in luciferase activity in GAL4-AUTS2 cells with or without a P300 inhibitor C646 for 16 h, before 24 h doxycycline induction. **d**, Mapping of AUTS2 domains required for transcriptional activation. Shown are schematic representations of stably expressed, truncated versions of AUTS2 fused to GAL4 (left) and their ability to stimulate luciferase activity (right). **e**, ChIP in stable cell lines expressing the candidates shown in **d** and performed with antibodies against GAL4 and P300.

GAL4–AUTS2 activate transcription dependent upon the ability to recruit P300 (Fig. 4e). Thus, through its recruitment of CK2 and interaction with P300, AUTS2 is the key to converting repressive PRC1 to a transcriptional activator.

AUTS2 regulates CNS gene expression

The tissue distribution of AUTS2 during development was assessed using immunohistochemistry (IHC) and validated AUTS2 antibody (Extended Data Fig. 5a). AUTS2 is highly expressed in the CNS of E15 mouse embryo, with the highest level being apparent in the neocortex (Extended Data Fig. 5b). AUTS2 was detected in postnatal day 30 (P30) adult mouse brain sections, especially in hippocampus, cerebral cortex, and Purkinje cells in the cerebellum, but at considerably lower levels (Extended Data Fig. 5c). Western blot analyses confirmed AUTS2 expression in the brain as a function of development (Extended Data Fig. 5d), consistent with reported *Auts2* mRNA levels¹⁹. Furthermore, AUTS2 is probably expressed primarily in neurons based on immunofluorescence analysis (Extended Data Fig. 5e) and prior studies¹⁹.

We next performed ChIP followed by deep sequencing (ChIP-seq) to determine the genomic localization of AUTS2 and RING1B, and the co-presence of various histone modifications and Pol II using P1 mouse brain. AUTS2 was found predominantly in the ± 5 kb region surrounding transcriptional start sites (TSS) (Extended Data Fig. 6a, b). Several promoters were co-bound by AUTS2, Pol II, and histone modifications associated with active transcription, including histone H3 acetylated at lysine 27 (H3K27ac), and H3K4me3 (Fig. 5a). Moreover, the gene bodies following these promoters were decorated with trimethylation of histone H3 at lysine 36 (H3K36me3), a modification linked with Pol II elongation (Fig. 5a). H3K27me3 was not detected at these AUTS2 bound promoters (Fig. 5a), consistent with the absence of detectable CBX protein in PRC1.5–AUTS2. Importantly, although RING1B was enriched at these promoters, its enzymatic product H2AK119ub1 was absent (Fig. 5a).

We then performed genome-wide analysis and identified significantly enriched peaks, as reported previously¹⁶. Approximately 50% of

AUTS2-containing peaks comprise regions bound by H3K27ac, whereas only 9.0% were bound by H2AK119ub1 (Fig. 5b), consistent with the intrinsic E3 ligase activity of RING1A/B being suppressed in PRC1.5–AUTS2, as shown *in vitro* (Fig. 3a). Amongst the AUTS2 target genes identified, ~35.2% comprise the top 25% highly transcribed genes in mouse brain based on reads per kilobase of exon per million reads mapped (RPKM) values obtained from our RNA-seq analysis in P1 mouse whole brain (Extended Data Fig. 6c). In contrast, only 8.9% comprise the bottom 25% genes with the lowest RPKM values (Extended Data Fig. 6c). We next identified genes co-targeted by AUTS2 and RING1B, and those by BMI1 and RING1B as control. Out of 4,168 AUTS2 target genes, 1,488 were also bound by RING1B; whereas out of 1,919 BMI1 target genes, 1,137 were also bound by RING1B. The average expression of AUTS2 or AUTS2/RING1B target genes was significantly higher than those of BMI1 or BMI1/RING1B (Fig. 5c). Of note, the overlap between AUTS2- and RING1B-targeted genes is relatively low, indicating that AUTS2 may be recruited to chromatin through PRC1-independent mechanism(s), perhaps involving other AUTS2 interacting candidates identified by our TAP analysis (Fig. 1a).

AUTS2/RING1B co-targeted loci comprised higher levels of H3K27ac and Pol II, and lower levels of H2AK119ub1 and H3K27me3, relative to BMI1/RING1B co-targets (Fig. 5d). Moreover, P300 was localized to AUTS2-targeted loci and its global occupancy was higher on loci targeted by AUTS2/RING1B than those by BMI1/RING1B, as evidenced by ChIP-seq in mouse brain (Extended Data Fig. 7). Gene Ontology (GO) analysis of AUTS2 targets in mouse brain revealed enrichment of functional terms including “gene expression”, “abnormality of the forebrain”, and “abnormality of the cerebrum” (Fig. 5e; Supplementary Table 2), indicating a role of AUTS2 in positively regulating the CNS transcriptional program.

RING1B was similarly associated with AUTS2 target genes in 293 T-REx cells, as shown by ChIP-seq (Extended Data Fig. 8). A substantial portion of HA–AUTS2 peaks (approximately 30%: 1,664 out of 5,563 peaks) were also bound by HA–RING1B (Extended Data Fig. 8c).

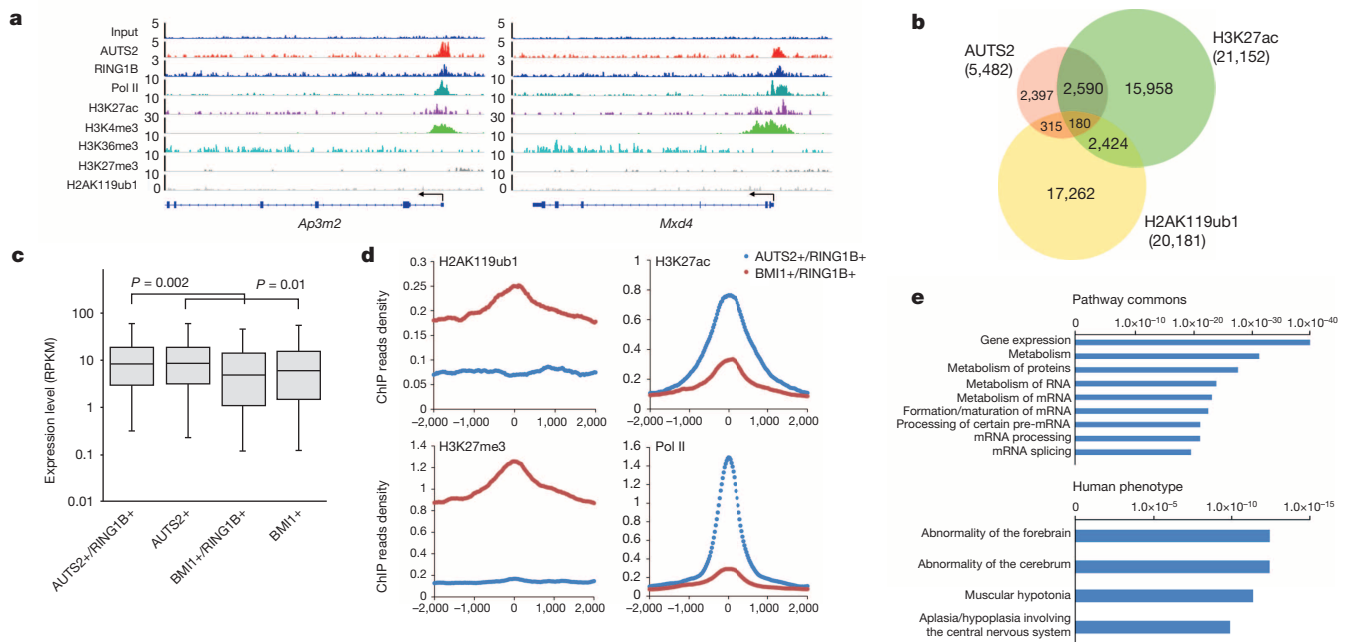


Figure 5 | Regulation of neuronal gene expression by AUTS2. **a**, IGV browser views for input, AUTS2, RING1B, Pol II, H3K27ac, H3K4me3, H3K36me3, H3K27me3, and H2AK119ub1. ChIP-seq performed in P1 mouse brain at two representative loci indicated. The y axis corresponds to ChIP-seq signal intensity. **b**, Venn diagram showing the overlap among target regions of AUTS2, H3K27ac, and H2AK119ub1. **c**, Expression analysis of genes targeted by AUTS2/RING1B, AUTS2 alone, BMI1/RING1B, and BMI1 alone.

RPKM values are obtained from our RNA-seq data in mouse whole brain. **d**, ChIP reads density plots for levels of indicated histone marks and Pol II at loci co-targeted by AUTS2/RING1B and BMI1/RING1B. A ± 2 kb window relative to the centre of peaks is shown. **e**, GO analysis of targeted genomic regions identified by AUTS2 ChIP-seq. The x axis (in logarithmic scale) corresponds to the binomial raw P values.

NFH-AUTS2 induction in 293 T-REx cells had no effect on *RING1B* and *RYBP* mRNA levels as assessed by reverse transcription qPCR (RT-qPCR), whereas mRNA levels of several genes among the top 20 HA-AUTS2 targets were generally increased (Extended Data Fig. 8f), consistent with PRC1.5-AUTS2 mediating transcriptional activation.

Phenotypes in *Auts2* KO mice

AUTS2 function in the brain was assessed using *Auts2* conditional knockout (cKO) mice generated by crossing mice harbouring *LoxP* sites flanking exon 7 of *Auts2* with mice carrying nestin promoter-driven Cre recombinase (Fig. 6a; Extended Data Fig. 9a, b). Although AUTS2 disruptions normally occur on one of the two alleles in patients, we characterized full homozygous knockout (KO) as well as heterozygous knockout of *Auts2* (Het) to better understand the *Auts2* phenotype, as well as understand gene dosage effects. In humans, ~80% of all AUTS2 disruptions are associated with either low birth weight or small stature²¹. Consistent with this observation, we observed both a striking visual (Fig. 6a) and quantitative (Fig. 6b; Extended Data Fig. 9c) reduction in the size of the *Auts2* knockout relative to wild-type littermates, with heterozygotes showing an intermediate phenotype across early development. Developmental delay typically encompasses impairments in reaching normal sensorimotor, cognition and communication (for example, speech) milestones, characteristics of the AUTS2 phenotype²¹. We explored such developmental milestones in mice as a function of *Auts2* disruption using a pre-weanling behavioural test battery including basic motor skills (for example, righting reflex) and ultrasonic vocalizations (USVs) following maternal separation, often impaired in a variety of mouse models of neurodevelopmental disorders^{24–26}. Across early development, knockout mice were deficient in both righting reflex and ultrasonic vocalizations emitted (Fig. 6c, d), as well as in negative geotaxis (Extended Data Fig. 9d). Although the knockouts showed a significantly smaller milk band at P1, early malnutrition or abnormal maternal behaviour may not be fully responsible for all of the observed phenotypes, given that the heterozygotes showed phenotypes similar to the knockout although having a milk band indistinguishable from wild type (Extended Data Fig. 9e). Although it is difficult to directly infer human pathology from mouse phenotypes, these results strongly suggest that a

gene dosage-dependent disruption of mouse *Auts2* resulted in impaired developmental phenotypes characteristic of AUTS2 disruption in humans.

Several selected genes co-targeted by AUTS2 and RING1B exhibited decreased mRNA levels in the knockout relative to wild-type mouse whole brains as evidenced by RT-qPCR (Extended Data Fig. 10). In particular, disruption of *Nat8l* has been associated with many of the phenotypes observed in AUTS2 patients (for example, growth delay, intellectual disability)^{27–29}, suggesting that *Auts2* disruption in these mice may lead to developmental defects through altered expression of its target genes.

Discussion

Active promoters have been linked previously with the presence of PRC2 or PRC1 (refs 30–33), yet the heterogeneous compositions of these complexes hampered further investigations. Here we identified AUTS2 as the component that renders PRC1 capable of transcription stimulation, through AUTS2-mediated recruitment of CK2 and P300 (Fig. 6e). These findings underscore that natural variation in the constituents of PRC1 complexes can lead to PRC1 adopting unexpected roles in coordinating specific cellular gene expression profiles. Our findings may set a precedent for other dynamic alterations in the regulatory properties of PRC1 and perhaps PRC2, based on their constituent components.

Despite much evidence from human genetics indicating that AUTS2 disruption is associated with neurological disorders including ASD, AUTS2 function was unclear²¹. Our results provide the first, to our knowledge, causal evidence for *Auts2* disruption leading to specific behavioural phenotypes associated with the human condition. Although the precise neurobiological mechanisms underlying these developmental phenotypes are yet to be elucidated, AUTS2 association with active transcription and H3K4me3 point to its having a key role in regulating early transcriptional programs associated with normal brain development^{34,35}. AUTS2 is highly enriched in the prefrontal cortex (PFC) and recent evidence suggests that the genome-wide distribution of H3K4me3 peaks fail to exhibit the normal developmental shift at a number of genes implicated in ASD pathology (for example, *PARK2*, *NLGN4Y*, *SHANK3*), in the PFC of ASD patients³⁵. Furthermore, recent evidence provide a strong case for aberrant epigenetic regulation in the cerebellum of ASD patients^{36,37}. These findings combined with the enrichment of AUTS2

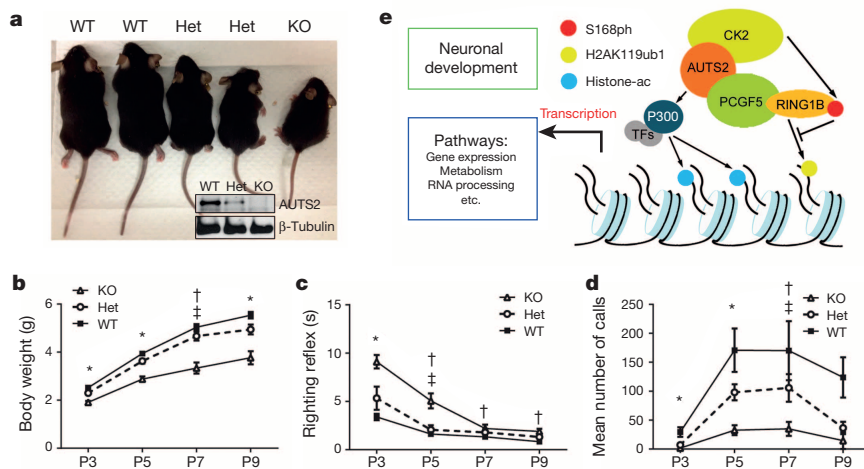


Figure 6 | Effect of the *Auts2* knockout on phenotypes associated with neurodevelopmental disorders. **a**, Comparison of the body size of wild-type, heterozygotes, and knockout littermates. Insert shows immunoblotting of respective whole brain extracts. **b**, Both knockout and heterozygote mice weigh less than wild-type littermates across P3–P9. Data reported in all behavioural figures are mean and error bars are standard error of the mean. Post-hoc difference ($P < 0.05$) is indicated by * (between all three genotypes), † (between wild type and knockout), and ‡ (between heterozygote and knockout). Total numbers of mice used in the behavioural analyses are as following (the range reflects different numbers used for each behavioural test): at P3, wild type $n = 11–17$, heterozygote $n = 3–10$, knockout $n = 6–11$; at P5, wild type

$n = 23–27$, heterozygote $n = 8–13$, knockout $n = 18–19$; at P7, wild type $n = 7–11$, heterozygote $n = 3$, knockout $n = 8$; at P9, wild type $n = 4–8$, heterozygote $n = 2–7$, knockout $n = 3$. Analysis was performed in at least three different litters of pups for each behavioural test. **c**, Knockout mice show impairment in righting reflex relative to wild type from P3–P9, whereas heterozygotes are not impaired. **d**, Both heterozygote and knockout mice show significantly less ultrasonic vocalizations (USVs) than wild type following maternal separation across the majority of developmental time points measured. **e**, A model for PRC1.5-AUTS2-mediated transcriptional activation. See main text for details.

in cerebellar Purkinje cells and the PFC suggest that these brain regions may be involved in conferring the *Auts2* phenotype²⁶. The novel role of AUTS2 in modulating PRC1 activity to effectively remove its repressive function and exploit the complex to attain activated transcription will probably encourage alternate directions in addressing the challenges of ASD and related neurological diseases.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 11 August; accepted 3 October 2014.

- Simon, J. A. & Kingston, R. E. Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Mol. Cell* **49**, 808–824 (2013).
- Müller, J. & Verrijzer, P. Biochemical mechanisms of gene regulation by polycomb group protein complexes. *Curr. Opin. Genet. Dev.* **19**, 150–158 (2009).
- Morey, L. & Helin, K. Polycomb group protein-mediated repression of transcription. *Trends Biochem. Sci.* **35**, 323–332 (2010).
- Schuettengruber, B. & Cavalli, G. Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice. *Development* **136**, 3531–3542 (2009).
- Jaenisch, R. & Young, R. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* **132**, 567–582 (2008).
- Sparmann, A. & van Lohuizen, M. Polycomb silencers control cell fate, development and cancer. *Nature Rev. Cancer* **6**, 846–856 (2006).
- Schwartz, Y. B. & Pirrotta, V. Polycomb silencing mechanisms and the management of genomic programmes. *Nature Rev. Genet.* **8**, 9–22 (2007).
- Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343–349 (2011).
- Müller, J. *et al.* Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. *Cell* **111**, 197–208 (2002).
- Czermin, B. *et al.* *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111**, 185–196 (2002).
- Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039–1043 (2002).
- Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* **16**, 2893–2905 (2002).
- Wang, H. *et al.* Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873–878 (2004).
- de Napolles, M. *et al.* Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7**, 663–676 (2004).
- Francis, N. J., Kingston, R. E. & Woodcock, C. L. Chromatin compaction by a polycomb group protein complex. *Science* **306**, 1574–1577 (2004).
- Gao, Z. *et al.* PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol. Cell* **45**, 344–356 (2012).
- Tavares, L. *et al.* RYBP-PRC1 complexes mediate H2A ubiquitylation at Polycomb target sites independently of PRC2 and H3K27me3. *Cell* **148**, 664–678 (2012).
- Morey, L., Aloia, L., Cozzuto, L., Benitah, S. A. & Di Croce, L. RYBP and Cbx7 define specific biological functions of Polycomb complexes in mouse embryonic stem cells. *Cell Rep.* **3**, 60–69 (2013).
- Bedogni, F. *et al.* Autism susceptibility candidate 2 (*Auts2*) encodes a nuclear protein expressed in developing brain regions implicated in autism neuropathology. *Gene Expr. Patterns* **10**, 9–15 (2010).
- Sultana, R. *et al.* Identification of a novel gene on chromosome 7q11.2 interrupted by a translocation breakpoint in a pair of autistic twins. *Genomics* **80**, 129–134 (2002).
- Oksenberg, N. & Ahituv, N. The role of *AUTS2* in neurodevelopment and human evolution. *Trends Genet.* **29**, 600–608 (2013).
- Beunders, G. *et al.* Exonic deletions in *AUTS2* cause a syndromic form of intellectual disability and suggest a critical role for the C terminus. *Am. J. Hum. Genet.* **92**, 210–220 (2013).
- Oksenberg, N., Stevison, L., Wall, J. D. & Ahituv, N. Function and regulation of *AUTS2*, a gene implicated in autism and human evolution. *PLoS Genet.* **9**, e1003221 (2013).
- Scattoni, M. L., Gandhi, S. U., Ricceri, L. & Crawley, J. N. Unusual repertoire of vocalizations in the BTBR T+tf/J mouse model of autism. *PLoS ONE* **3**, e3067 (2008).
- Scattoni, M. L., Crawley, J. & Ricceri, L. Ultrasonic vocalizations: a tool for behavioural phenotyping of mouse models of neurodevelopmental disorders. *Neurosci. Biobehav. Rev.* **33**, 508–515 (2009).
- Fujita, E. *et al.* Ultrasonic vocalization impairment of Foxp2 (R552H) knockin mice related to speech-language disorder and abnormality of Purkinje cells. *Proc. Natl Acad. Sci. USA* **105**, 3117–3122 (2008).
- Wiame, E. *et al.* Molecular identification of aspartate N-acetyltransferase and its mutation in hypocacetylaspartia. *Biochem. J.* **425**, 127–136 (2010).
- Boltshauser, E. *et al.* Follow-up of a child with hypocacetylaspartia. *Neuropediatrics* **35**, 255–258 (2004).
- Furukawa-Hibi, Y. *et al.* Absence of SHAT1/Nat8l reduces social interaction in mice. *Neurosci. Lett.* **526**, 79–84 (2012).
- Kondo, T. *et al.* Polycomb potentiates *Meis2* activation in midbrain by mediating interaction of the promoter with a tissue-specific enhancer. *Dev. Cell* **28**, 94–101 (2014).
- Mousavi, K., Zare, H., Wang, A. H. & Sartorelli, V. Polycomb protein Ezh1 promotes RNA polymerase II elongation. *Mol. Cell* **45**, 255–262 (2012).
- Yu, M. *et al.* Direct recruitment of Polycomb repressive complex 1 to chromatin by core binding transcription factors. *Mol. Cell* **45**, 330–343 (2012).
- Frangini, A. *et al.* The Aurora B kinase and the Polycomb protein Ring1B combine to regulate active promoters in quiescent lymphocytes. *Mol. Cell* **51**, 647–661 (2013).
- Cheung, I. *et al.* Developmental regulation and individual differences of neuronal H3K4me3 epigenomes in the prefrontal cortex. *Proc. Natl Acad. Sci. USA* **107**, 8824–8829 (2010).
- Shulha, H. P., Cheung, I., Guo, Y., Akbarian, S. & Weng, Z. Coordinated cell type-specific epigenetic remodeling in prefrontal cortex begins before birth and continues into early adulthood. *PLoS Genet.* **9**, e1003433 (2013).
- James, S. J., Shpileva, S., Melnyk, S., Pavliv, O. & Pogribny, I. P. Complex epigenetic regulation of engrailed-2 (*EN-2*) homeobox gene in the autism cerebellum. *Transl. Psychiatry* **3**, e232 (2013).
- Ladd-Acosta, C. *et al.* Common DNA methylation alterations in multiple brain regions in autism. *Mol. Psychiatry* **19**, 862–871 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank S. Zhang, R. Bonasio, S. Shen, and L. Cohen for assistance in analysis with genome-wide studies; H. Zheng for mass spectrometry analysis; L. Chiriboga, C. Loomis and the Experimental Pathology Cores at NYU Langone Medical Center for assistance in immunohistochemical analysis; D. Hernandez for assistance in various experiments. We appreciate the help of C. Guo and the transgenic facility at Howard Hughes Medical Institute's (HHMI) Janelia Farm Research Campus in generating the *Auts2* cKO mice. We are grateful to V. Bardwell and L. Di Croce for providing us PCGF1 and RING1B antibodies, respectively. We also thank A. Tarakhovskiy for his generous help to our mouse work after Hurricane Sandy and L. Vales for valuable discussions and extensive reading of the manuscript. This work was supported as follows: the biochemical analysis of AUTS2 by the National Institute of Health grant GM-64844, and the mouse work by a Pilot Award from the Simons Foundation Autism Research Initiative (SFARI). A.S. is supported by grants from NIH (1DP2MH100012-01), Seaver Autism Foundation, and The Brain and Behavioral Research Fund Young Investigator Award (number 18194). Z.G. was supported by the SFARI pilot award and by an NIH training grant (5T32CA160002). P.L. was supported by an NIH postdoctoral fellowship (1F32GM105275). J.M.S. was supported by an NIH postdoctoral fellowship (F32AA022842) as well as by the Simons Foundation as a Junior Fellow.

Author Contributions Z.G., P.L., J.M.S. and M.v.S. performed experiments. All authors contributed to data analysis, experimental design and manuscript writing.

Author Information All ChIP-seq and RNA-seq data has been deposited into GEO database with the accession number GSE60411. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.R. (danny.reinberg@nyumc.org).

METHODS

Plasmids. For construction of pINTO-NFH-AUTS2, cDNA for human AUTS2 was purchased from ATCC and cloned into the vector pINTO-NFH carrying N-terminal Flag and HA tags, as described previously¹⁶. For pINTO-GAL4-AUTS2, the same cDNA was cloned into pINTO-GAL4 with a GAL4 DNA binding sequence replacing the N-terminal Flag and HA tags. For truncated GAL4-AUTS2, PCR was used to amplify the fragments of interest before inversion to the pINTO-GAL4 vector. For RING1B, PCGF4, and PCGF5, cDNA obtained from digestion of pINTO-NFH-RING1B, pINTO-NFH-PCGF4, and pINTO-NFH-PCGF5, respectively, were cloned into pINTO-GAL4. For protein expression in insect cells, cDNAs of interest were cloned into pFastbac vectors.

Cell culture. Stable cells expressing pINTO-NFH-AUTS2 were generated from human 293 T-REx cells and maintained as described previously¹⁶. For those expressing GAL4-fusion proteins, vector control (pINTO-GAL4) or constructs made in this vector were transfected into 293 T-REx-luciferase cells containing a stable integrated 5XGal4RE-tk-Luc-neo construct³⁸, and selected with 200 $\mu\text{g ml}^{-1}$ zeocin, 10 $\mu\text{g ml}^{-1}$ blasticidin, and 400 $\mu\text{g ml}^{-1}$ G418.

RNA interference. Cells were transfected with siRNAs using Lipofectamine RNAiMAX (Life Technologies) according to the manufacturer's protocol. All human siRNAs were purchased: si-ctrl (AllStars Negative Control siRNA, cat no. SI03650318) and si-RING1B (cat no. SI00095543, target sequence: 5'-TGGGCTAGAGCTTG ATAATAA-3') were from Qiagen; si-P300 (cat no. 4392420, ID no. s4695, target sequence: 5'-GGACTACCCTATCAAGTAA-3') was from Ambion; si-PCGF5 (cat no. L-007089-00-0005, target sequences: 5'-CAACAACAGTACGGAATG-3', 5'-GAGGTTGGACA ATACATTA-3', 5'-ACAAATTGCTATCTGTCTA-3', 5'-G AAGAAATTCATTCGATGT-3') and si-AUTS2 (cat# L-013932-00-0005, target sequences: 5'-TGACAGAGATAGAGATGTA-3', 5'-AGACTCATCTGTTAGT AAA-3', 5'-GAAAGGCTCAGTGATAGTT-3', 5'-CACATAA GCTGGACTTT GG-3') were from Thermo Scientific.

Affinity purification, protein identification, and glycerol gradient. To identify proteins associated with AUTS2, tandem affinity purification was performed in nuclear extract (NE) of 293 T-REx cells expressing NFH-AUTS2, as described previously¹⁶. Briefly, nuclear extract prepared from 30 \times 150 mm plates of cells was incubated with Flag M2 beads at 4 °C overnight. After 5 washes, proteins bound on the M2 beads were eluted with 500 μl Flag peptides at 250 $\mu\text{g ml}^{-1}$ at 4 °C for 1 h. The M2 eluate was then incubated with 30 μl HA beads at 4 °C for 4 h. The HA beads were washed 5 times and proteins eluted with 100 μl glycine (0.1 M, pH 2.0), and then neutralized by adding 6.5 μl Tris solution (1.5 M, pH 8.8), resulting in the final HA eluate, which was then analysed by LC-MS.

For glycerol gradient analysis, 100 μl M2 eluate was subjected to ultracentrifugation followed by fractionation, as described previously¹⁶. The odd-numbered fractions were then subjected to western blotting.

Immunoprecipitation and *in vitro* interaction assay. Immunoprecipitation experiments were performed as described previously¹⁶ with certain modifications. For immunoprecipitation, nuclear extract prepared from the indicated sources was incubated with antibodies before addition of protein G beads, or with Flag or HA beads at 4 °C for 3 h. For *in vitro* interaction assay, 60 h after infection with baculovirus for proteins of interest, Sf9 cells were harvested lysed in lysis buffer (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 4 mM MgCl₂, 0.4 mM EDTA, 2 mM DTT, 20 mM β -glycerophosphate, 20% glycerol, 0.4 mM PMSF, 1 $\mu\text{g ml}^{-1}$ pepstatin A, 1 $\mu\text{g ml}^{-1}$ leupeptin, 1 $\mu\text{g ml}^{-1}$ aprotinin), followed by centrifugation. Lysate containing the indicated proteins were incubated with Flag or HA beads at 4 °C for 3 h. Beads were then washed 5 times and eluted with 100 μl glycine (0.1 M, pH 2.0), and then neutralized by adding 6.5 μl Tris solution (1.5 M, pH 8.8). The eluted samples were mixed with SDS sample buffer and analysed by SDS-PAGE, followed by immunoblotting.

Luciferase reporter assay. 293T-REx-luciferase cells³⁸ stably transfected with pINTO-GAL4 vector control or with inserts of interest were treated with 100 ng ml⁻¹ doxycycline. Cells were lysed by adding 250 μl of ice-cold lysis buffer (0.2% Triton X-100, 100 mM potassium phosphate, pH 7.8, and 1 mM DTT) and shaking for 10 min at 4 °C. The cell lysate was centrifuged at 20,000g for 10 min and the protein concentration of the resulting supernatant was determined by Bradford assay. 30 μg of the supernatant was assayed for luciferase activity using luciferase assay substrate (Promega).

Purification of crosslinked nuclei from whole mouse brains. Nuclei from whole mouse brains were purified as previously described for cerebellar mouse nuclei³⁹ including minor changes. Briefly, CD1 pups at P1 were isolated from the embryonic sac and their whole brains quickly dissected. The whole brains were homogenized with a glass douncing homogenizer using first a loose, then tight pestle (Kimble Chase; 1984-10002). The cell homogenate was fixed with a final concentration of 1% paraformaldehyde for 8 min at room temperature and the reaction quenched with 0.125 M glycine for 5 min at room temperature. To isolate the nuclei, the fixed homogenate was spun through a 29% iodixanol cushion and the nuclei pellet resuspended in resuspension buffer (0.25 M sucrose, 25 mM KCl, 5 mM

MgCl₂, 20 mM tricine, pH 7.8 and 10% donkey serum) supplemented with 0.15 mM spermine, 0.5 mM spermidine and EDTA-free protease inhibitor cocktail (Roche; 11836170001). Resuspended nuclei were counted using a hemocytometer and aliquots of 10⁸ nuclei pelleted by centrifugation at 2,000g for 15 min at 4 °C. Nuclei pellets were flash-frozen in liquid nitrogen and stored at -80 °C until further analysis.

ChIP and ChIP-seq. Chromatin IP (ChIP) was performed as described previously¹⁶. Briefly, crosslinked and isolated nuclei were sonicated using a Diagenode Bioruptor to an average size of ~250 bp. After pre-clearing with BSA-blocked protein G Sepharose, chromatin was incubated with antibodies at 4 °C overnight. The chromatin immunocomplexes were recovered with the same BSA-blocked protein G beads. For ChIP-seq library construction, 5–10 ng of DNA extracted from the chromatin immunocomplexes as described previously¹⁵. Libraries were prepared according to manufacturer's instructions (Illumina) and as described¹⁵. Immunoprecipitated DNA was first end-repaired using End-It Repair Kit (Epicentre), tailed with an A using Klenow exo minus (NEB M0212), and ligated to custom adapters with LigaFast (Promega, no. M8225). Fragments of 350 \pm 50 bp were size-selected and subjected to ligation-mediated PCR amplification (LM-PCR), using Phusion DNA polymerase (NEB M0530). Libraries were quantified with quantitative PCR using primers annealing to the adaptor sequence and sequenced at a concentration of 7 pM on an Illumina HiSeq 2000. All sequencing data has been deposited into GEO/NCBI with the accession number GSE60411.

ChIP-seq Analysis. ChIP-seq analysis was performed as described previously¹⁶ with certain modifications. Sequenced reads (36 bp) were aligned to the mouse reference genome (assembly mm9) for each ChIP-seq experiment in mouse brain, and to human reference genome (assembly hg19) for each ChIP-seq experiment in 293T-REx cells using Bowtie⁴⁰. Duplicated reads were removed with samtools⁴¹. ChIP-seq read density files were generated using igvtools and were viewed in Integrative Genomics Viewer (IGV)⁴². Significantly ($P < 0.01$) enriched peaks for each ChIP-seq data set were identified with QSESeq⁴³ and ranked according to the total number of reads mapping to them. Venn diagram of overlap among peaks (Fig. 5b; Extended Data Fig. 8c–e) was computed using the R statistical software package (<http://cran.r-project.org>). Correlation of AUTS2 target genes with their expression level (Extended Data Fig. 6c) was obtained by intersecting with a gene list ranked by RPKM values obtained by our RNA-seq study in P1 mouse brain. The same RPKM values were used to produce the box-and-whisker plot (Fig. 5c). ChIP reads density plots (Fig. 5d) were made using HOMER by calculating average tag densities across \pm 2 kb regions surrounding indicated peaks. Gene associated region annotations (Fig. 5e) were obtained with genomic regions enrichment of annotations tool (GREAT)⁴⁴. Genomic distribution of peaks relative to TSS (Extended Data Figs 6b, 8b) was obtained via HOMER.

***In vitro* H2A ubiquitination assays.** Assays were performed as described previously¹⁶ with certain modifications. Briefly, in the presence of 100 nM E1 (Boston Biochem), 500 nM UbcH5c (Boston Biochem), 10 μM HA-ubiquitin (Boston Biochem), 0.5 mM ATP, 0.1 $\mu\text{g ml}^{-1}$ creatine kinase, 25 mM phosphocreatine, 1 $\mu\text{g ml}^{-1}$ BSA, and 2 μl of 10 \times ubiquitination reaction buffer (500 mM Tris-HCl, pH 7.5, 50 mM MgCl₂, 10 mM DTT), reactions were assembled with reconstituted oligonucleosomes (~2.5 μg) and purified complexes or proteins in a total volume of 20 μl . After 1 h incubation at 37 °C, the reactions were stopped by boiling in SDS sample buffer, and then resolved on SDS-PAGE, followed by immunoblotting. To test the effect of CK2 (NEB, P6010) on ubiquitination (Figs 3a and b), recombinant RING1B-PCGF5-AUTS2 complex purified from Sf9 cells was incubated with CK2 in reaction buffer (20 mM Tris-HCl, pH 7.5, 50 mM KCl, and 10 mM MgCl₂) in the presence of ATP, creatine kinase, phosphocreatine, and BSA at 4 °C for overnight to minimize the inactivation of CK2. Other components were then added and incubated for additional 1 h at 37 °C. For the experiment with E1 pre-charged with ubiquitin (Extended Data Fig. 2c), 5-fold E1 was incubated with HA-ub and ATP at room temperature for 30 min and then quenched with 20 mM EDTA at room temperature for 10 min. All the other components except for nucleosomes were incubated at 4 °C for overnight and also quenched by EDTA. These two mixtures were then mixed and supplemented with nucleosomes and incubated at 37 °C for another 30 min.

***In vitro* CK2 kinase assays.** The reactions were assembled with components used in the *in vitro* H2A ubiquitination assays at the same concentrations, with the exception of 50 μM in the case of ATP. CK2 was then added at 300 nM along with 2 μCi [γ -³²P]ATP (PerkinElmer, BLU502H500UC), followed by incubation at 37 °C for 30 min. The reactions were then resolved on SDS-PAGE and visualized by radiograph. For identification of phosphorylation sites on CK2 substrates, [γ -³²P]ATP was omitted in the assays. After SDS-PAGE and Coomassie blue staining, bands corresponding to specific proteins were excised and sent for mass spectrometry analysis.

Mice. Mice harbouring *LoxP* sites flanking *Auts2* exon 7 were generated at the Janelia Farms Gene Targeting and Transgenics Resource Center and were backcrossed onto the C57Bl/6J strain. Nervous system-specific deletions of *Auts2* was performed by crossing to nestin promoter driven Cre mice purchased from the

Jackson Laboratory (stock no: 003771). After weaning, mice were housed with same sex littermates and had free access to lab chow and water. Subjects were maintained on a 12-h light/dark cycle (lights on at 06:00). The laboratory temperature remained at $21 \pm 1^\circ\text{C}$. All behavioural experiments were started at $09:00 \pm 1\text{ h}$ and performed under protocols approved by the NYU School of Medicine IACUC and in accordance with NIH Principles of Laboratory Animal Care guidelines.

Immunofluorescence analysis of mouse brain sections. Mouse brains were harvested at the designated time points and fixed in 4% paraformaldehyde overnight at 4°C . The brains were then subjected to step-wise alcohol dehydration followed by tissue clearing with Histo-Clear (National Diagnostics) and embedded in paraffin (Fisher Scientific). Brains were sectioned to $16\text{ }\mu\text{m}$ and immunofluorescence was performed using the following antibodies: rabbit anti-AUTS2 (1:50, lab made), mouse anti-NeuN (1:100, Millipore, clone A60), chicken anti-GFAP (1:2000, Abcam). Slides were mounted with SlowFade (Invitrogen) and imaged using a Leica microscope.

Developmental phenotyping. For the testing battery experiments mouse pups from multiple litters were used for the behavioural analysis. In this way, each data-point represents multiple replicates (litters). The target sample size for each genotype was 8–10 as this has previously been sufficient to detect statistical differences in these behaviours, however the sample size varied due to a number of uncontrollable factors such as litter composition, subjects' health, environmental confounds and statistical outliers (over 2 standard deviations from the mean). Behaviour testing order was randomized for each litter of pups. Investigators were blinded to mouse genotype during behavioural testing and analysis except when the genotype was readily apparent in the phenotype (for example, smaller stature). Total numbers of WT, Het and KO mice used in the behavioural analyses are as following (the range reflects different numbers used for each behavioural test): at P3, WT 11–17, Het 3–10, KO 6–11; at P5, WT 23–27, Het 8–13, KO 18–19; at P7, WT 7–11, Het 3, KO 8; at P9 WT 4–8, Het 2–7, KO 3. Mice harbouring Cre recombinase alone did not differ from other WT littermates on any phenotype assayed thus were pooled in all analyses. Male and female mice were pooled for all analyses as no genotype X sex interaction was observed for any behaviour measured. Behavioural tests were performed in the following order; ultrasonic vocalizations, righting reflex, negative geotaxis. Experimenters were blinded to *Auts2* genotype.

Ultrasonic vocalizations. Measurement of ultrasonic vocalizations (USVs) following maternal separation was performed as reported previously^{27,45}. Briefly, at post-natal day 3, pups were isolated from their mothers. Following isolation, pups were placed in a plastic container (7 cm diameter) with fresh bedding contained in a sound attenuating styrofoam box. A condenser microphone (CM16/CMPA; Avisoft)

descended $\sim 15\text{ cm}$ above the pup through a hole in the top of the styrofoam box. Vocalizations were recorded with the Avisoft Recorder software (Version 4.2) using the UltraSoundGate 116 Hb interface. Standard USV recording parameters were used²⁵. Each recording session lasted 3 min. The room was maintained at $21 \pm 1^\circ\text{C}$ during all procedures. Following generation of spectrograms, the number of USVs were counted by a trained experimenter blinded to experimental conditions with the Avisoft-SASLab Pro software.

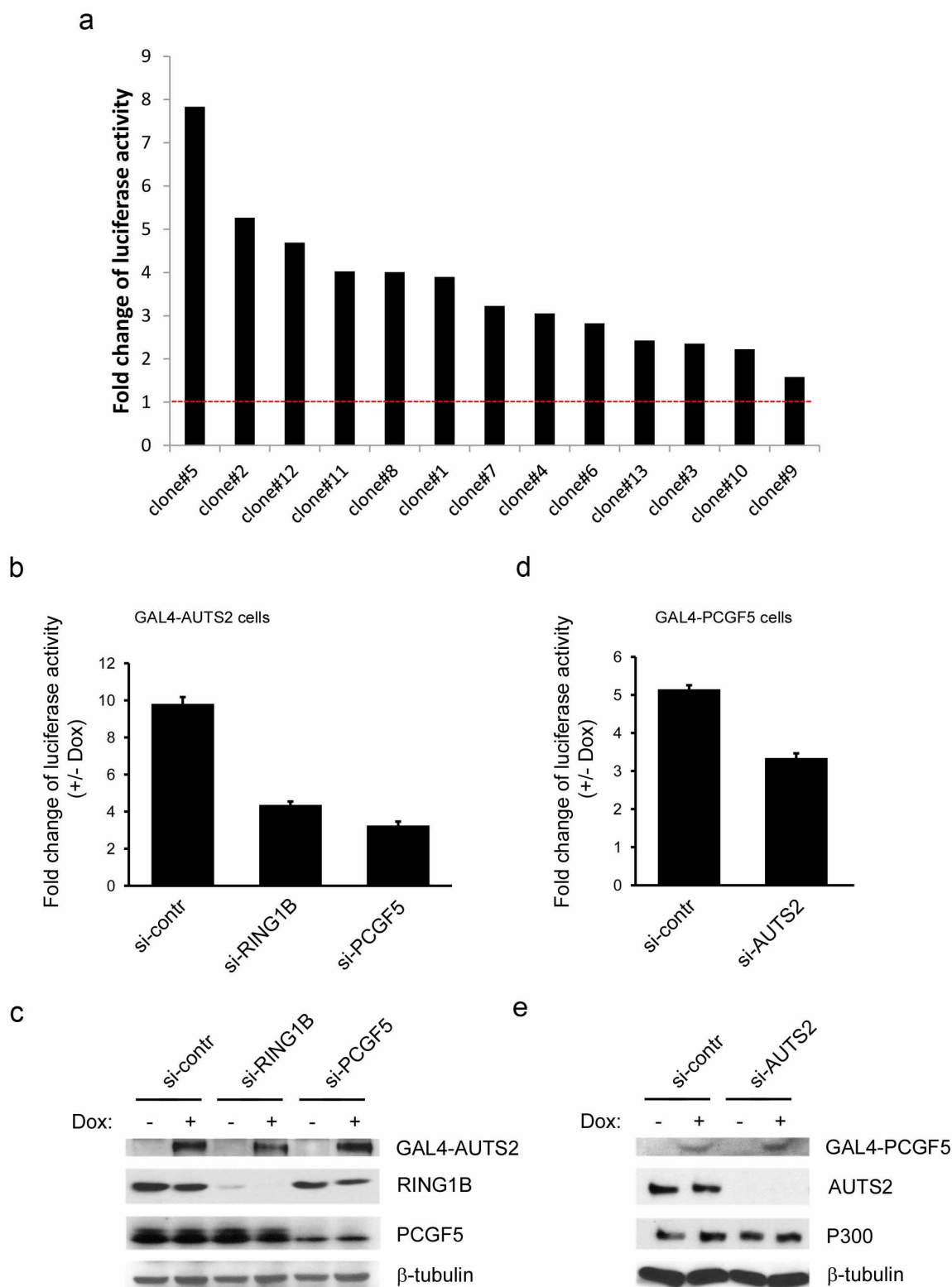
Righting reflex. Righting reflex was measured by recording the latency for a pup to place all four paws on the ground after being placed on its back. Each mouse was turned over twice with average latency to right itself take between both trials.

Negative geotaxis. Each pup was placed on a mesh grid angled at 45° with its nose facing towards the ground. Time to turn 180° and face upwards was measured. Maximum time allowed before rescue was 30 s on this task.

Milk band measure. At P1 the milk band was measured by visual inspection based on a standard sized milkband in a wild-type C57 P1 mouse pup. Milk in the stomach was judged as either empty, one-quarter, one-half or completely full by two independent observers with the consensus measure being used for final analysis.

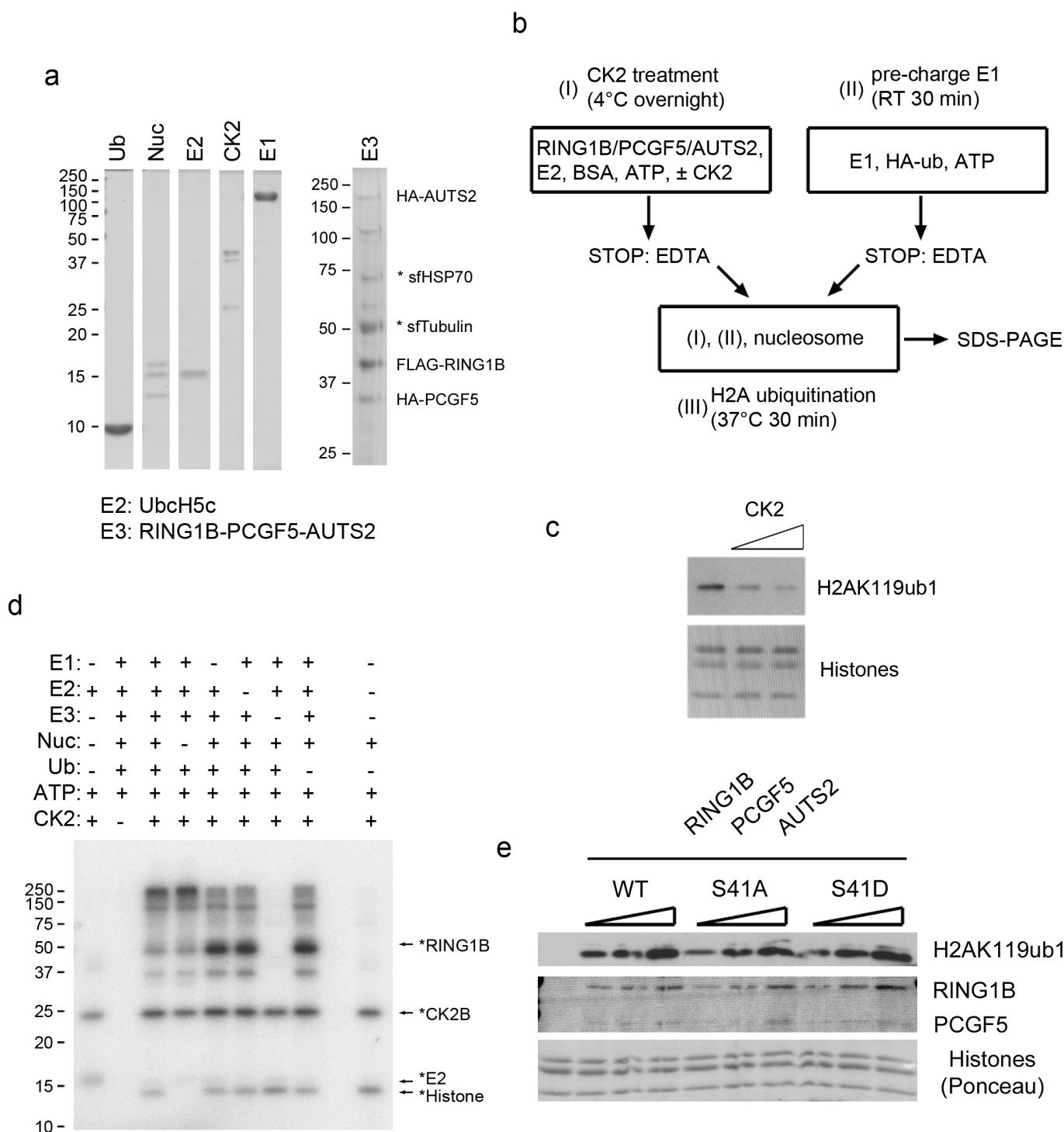
Statistics. At each developmental time point, a one-way analysis of variance (ANOVA) was performed to examine the effect of genotype on each phenotype measured. Fisher's LSD was used to test hypothesized differences between genotypes. Although no power analysis was performed *a priori*, sample sizes were determined based on prior studies showing samples sizes necessary to achieve significance (see above). Experimental α was set ≤ 0.05 .

38. Vaquero, A. *et al.* Human SirT1 interacts with histone H1 and promotes formation of facultative heterochromatin. *Mol. Cell* **16**, 93–105 (2004).
39. Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929–930 (2009).
40. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
41. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
42. Robinson, J. T. *et al.* Integrative genomics viewer. *Nature Biotechnol.* **29**, 24–26 (2011).
43. Asp, P. *et al.* Genome-wide remodeling of the epigenetic landscape during myogenic differentiation. *Proc. Natl Acad. Sci. USA* **108**, E149–E158 (2011).
44. McLean, C. Y. *et al.* GREAT improves functional interpretation of *cis*-regulatory regions. *Nature Biotechnol.* **28**, 495–501 (2010).
45. Hofer, M. A., Shair, H. N. & Brunelli, S. A. Ultrasonic vocalizations in rat and mouse pups. *Curr. Protocols Neurosci.* <http://dx.doi.org/10.1002/0471142301.ns0814s17> (2002).



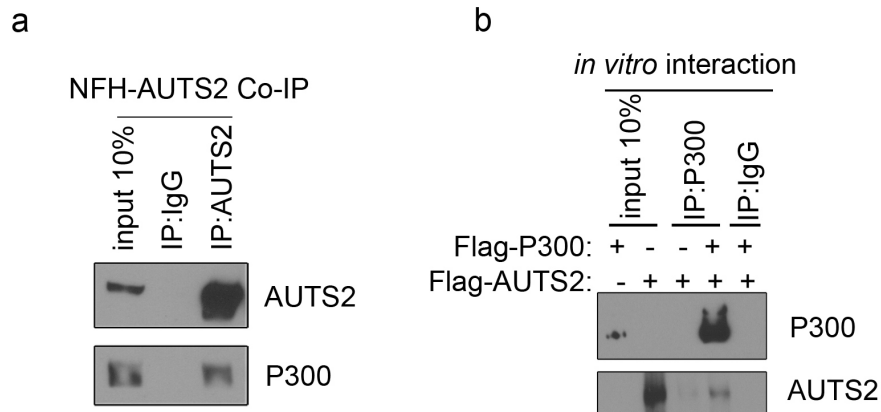
Extended Data Figure 1 | Requirement of the integrity of the PRC1-AUTS2 complex for transcriptional activation. **a**, Luciferase activity in screened stable cell clones expressing GAL4-PCGF5, 24 h after induction by doxycycline at $100 \mu\text{g ml}^{-1}$. **b**, Fold change in luciferase activity in GAL4-AUTS2 cells upon knockdown of RING1B or PCGF5. Cells were transfected with Lipofectamine 2000 RNAiMAX and siRNAs against RING1B or PCGF5, or control siRNAs for two days and then $100 \mu\text{g ml}^{-1}$ doxycycline was added to

induce GAL4-AUTS2 expression. Then 24 h after induction, luciferase activity was measured. Each value is the mean of three independent measurements. Error bars represent standard error. **c**, Immunoblotting of samples used for luciferase activity reporter assay as in **b** using the antibodies indicated. **d**, Fold change in luciferase activity in GAL4-PCGF5 cells upon knockdown of AUTS2. Cells were treated as in **b**. **e**, Immunoblotting of samples used for luciferase activity reporter assay as in **d** using the antibodies indicated.



Extended Data Figure 2 | H2A monoubiquitination assay and CK2 kinase assay performed with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$. **a**, Coomassie blue staining of factors used. **b**, Scheme for H2A monoubiquitination assay with E1 that was pre-charged with HA-ubiquitin (see Methods for details). **c**, Immunoblotting of H2A monoubiquitination assay as described in **b** with increasing amounts of CK2. **d**, Radiograph of CK2 kinase assay reaction products. The assay was assembled with the factors indicated, each at the same amount used in the H2A monoubiquitination assay (Methods). After incubation at 37 °C for 30 min,

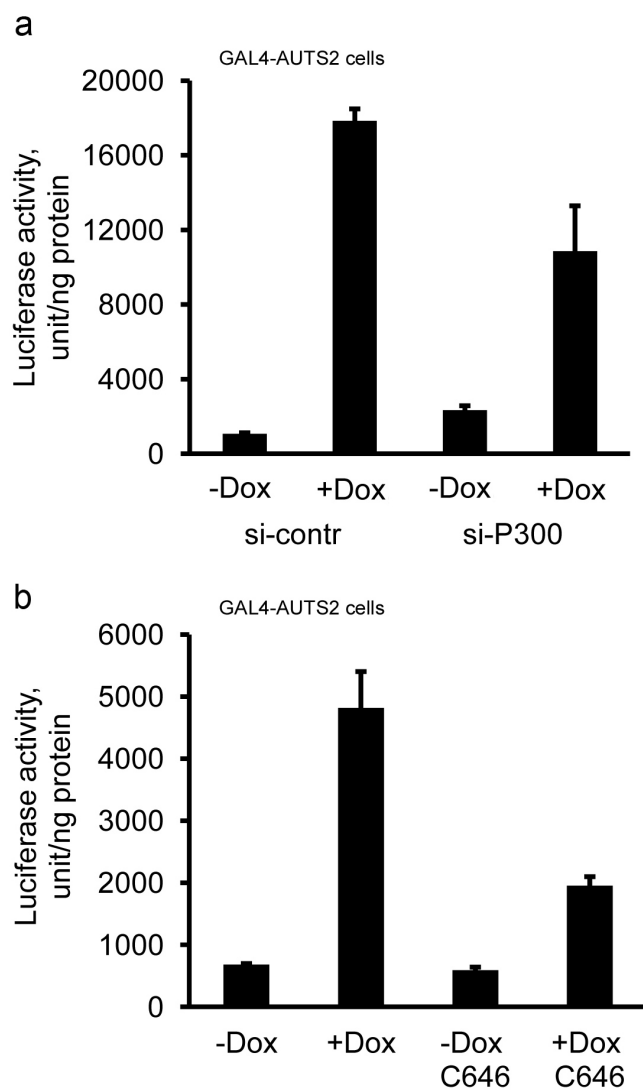
the assay was stopped by boiling in SDS loading buffer and resolved on SDS-PAGE. Besides CK2B, which was radio-labelled presumably due to autophosphorylation, phosphorylation of RING1B and PCGF5 was detected together with a species, indicated as *histone, dependent on the presence of nucleosomes. **e**, H2A monoubiquitination assay performed as in Fig. 3c, using increasing amounts of RING1B-PCGF5-AUTS2 containing either RING1B(S41A) (S41 to alanine), or RING1B(S41D) (S41 to aspartic acid), purified from Sf9 cells.



Extended Data Figure 3 | Interaction of AUTS2 and P300.

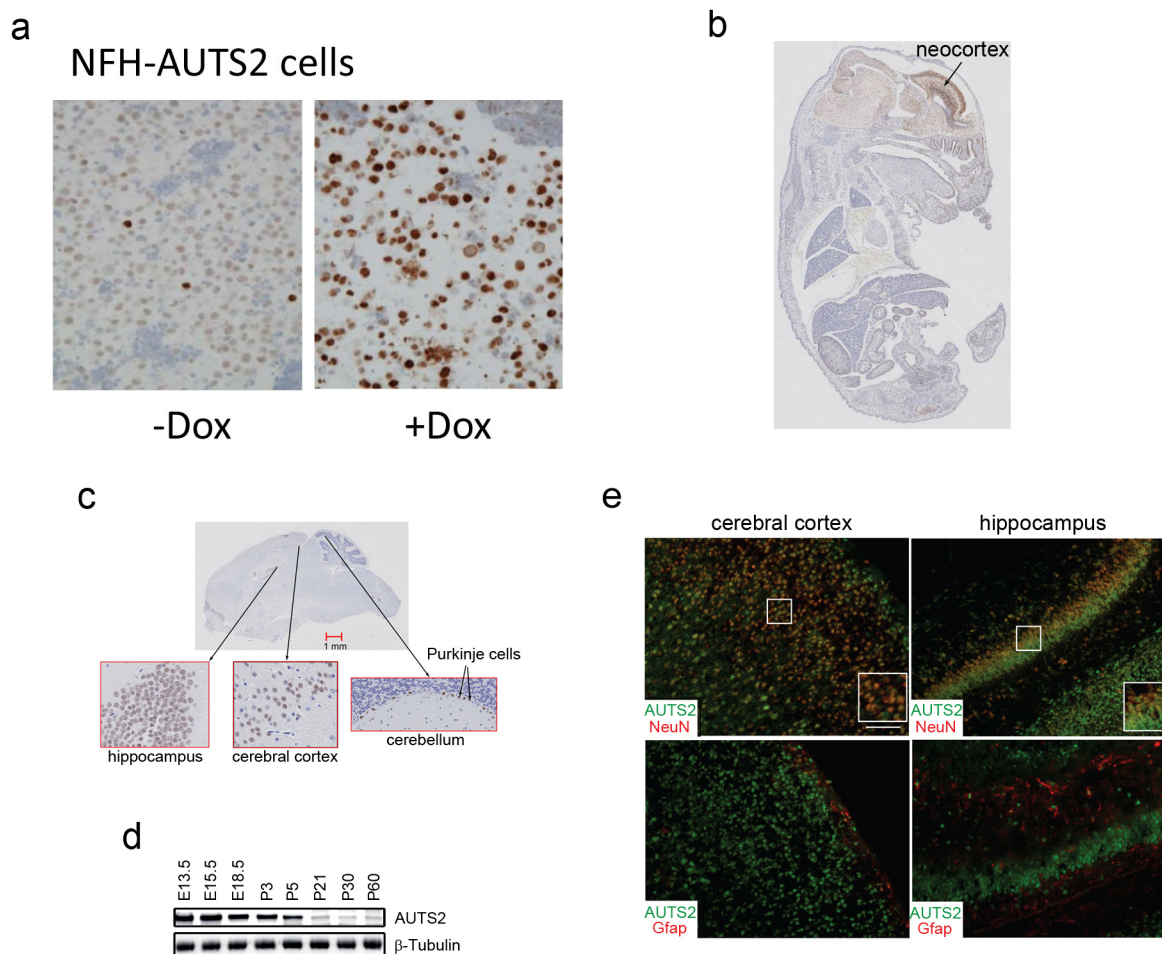
a, Immunoprecipitation from nuclear extract of 293T cells expressing NFH-AUTS2 using AUTS2 antibody, followed by western blotting for the

antigens indicated. **b**, Immunoprecipitation using recombinant proteins of P300 and AUTS2 purified from Sf9 cells and a P300 antibody, followed by western blotting using antibodies against P300 and AUTS2.



Extended Data Figure 4 | Luciferase activity without normalization.

a, b, Analysis using data from Fig. 4b, c, respectively.

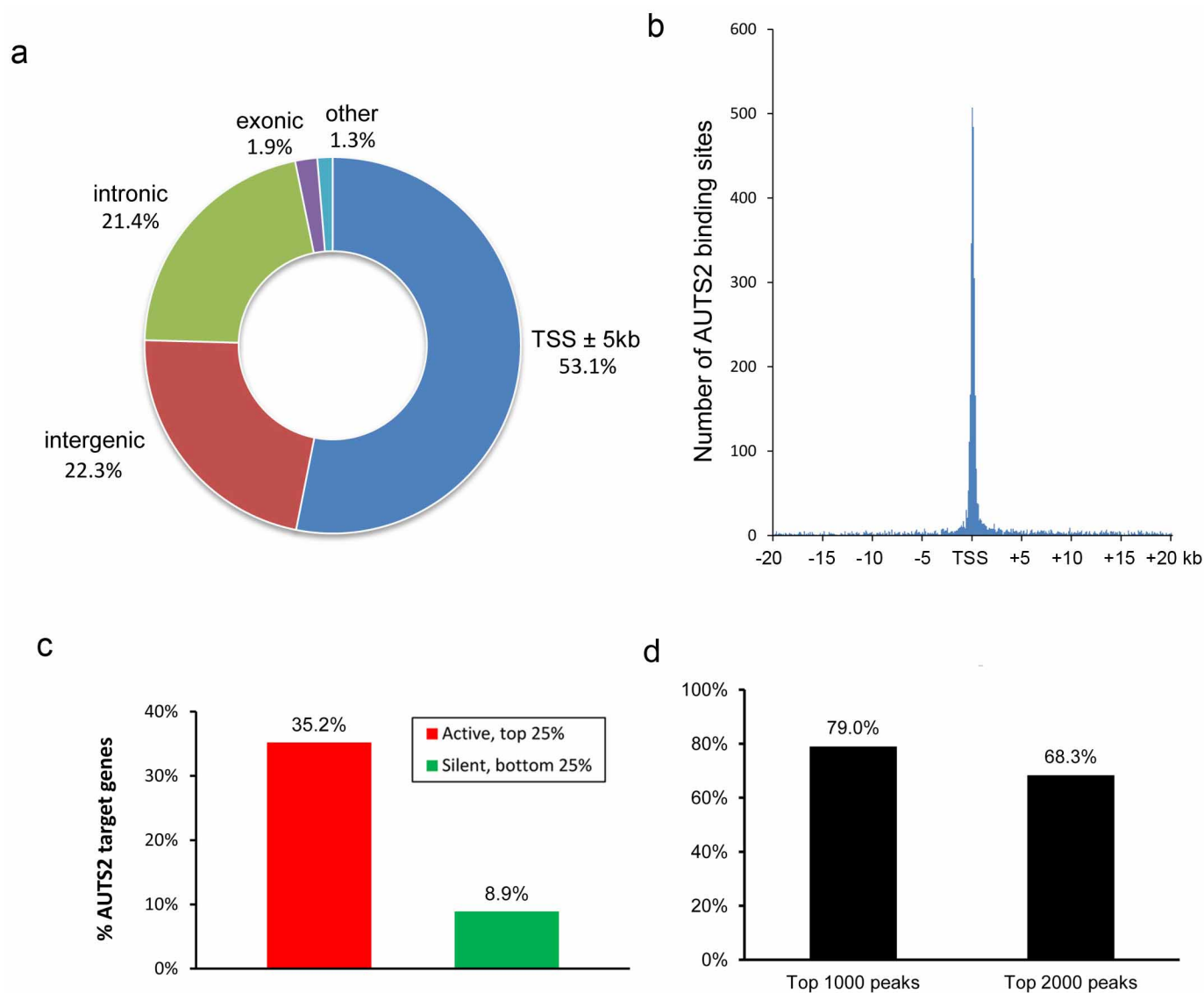


Extended Data Figure 5 | Expression of AUTS2 in mouse brain.

a, Validation of AUTS2 antibody by immunohistochemistry (IHC) in NFH-AUTS2 stable cells. Upon doxycycline induction, a stronger nuclear staining was detected compared with non-induction control, confirming the antibody we raised is suitable for IHC. **b**, Detection of AUTS2 protein in a mouse embryo at E15 by IHC with AUTS2 antibody. **c**, IHC analysis of a sagittal brain section from an adult mouse using AUTS2 antibody.

d, Expression of AUTS2 in the mouse brain. Immunoblotting was performed with whole brain extracts at various developmental stages as indicated.

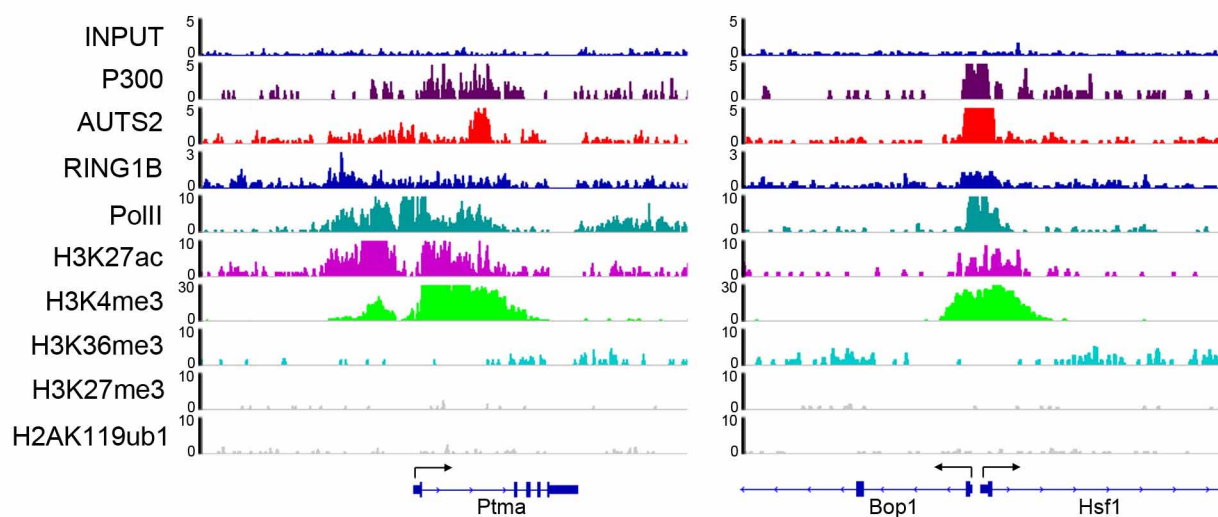
e, Immunofluorescence staining of AUTS2 in P3 mouse brain. AUTS2 expression is confined to neurons (top panels) as seen by co-localization with the neuronal marker NeuN in the cortex and hippocampus. AUTS2 does not co-localize with the glial marker GFAP (bottom panels) in the same regions.



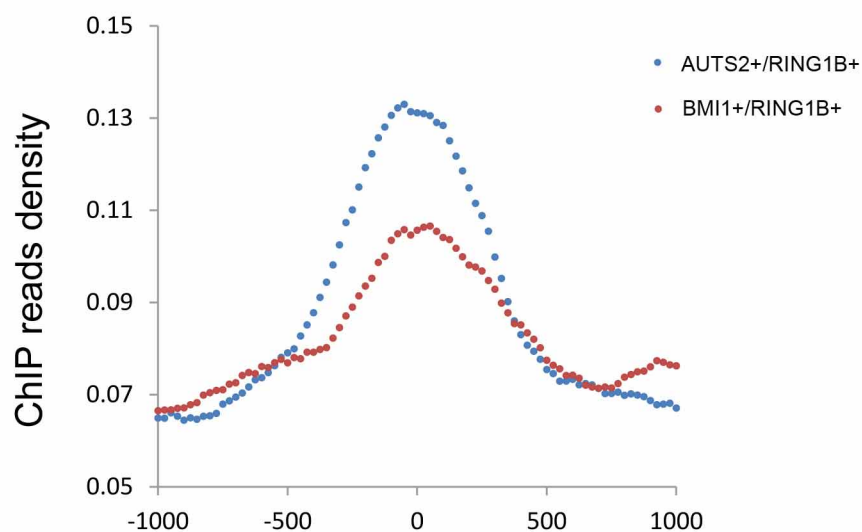
Extended Data Figure 6 | Genome-wide analysis of AUTS2 ChIP-seq signals. **a**, HOMER was used to compute the genomic distribution of AUTS2 peaks obtained from AUTS2 ChIP-seq in mouse brain. **b**, Histogram of the distribution of AUTS2 peaks relative to TSS, calculated via HOMER.

c, Percentage of AUTS2 target genes overlapped with highest (top 25%, red bar) and lowest (bottom 25%, green bar) expression levels of all genes in mouse brain. **d**, Percentage of overlapped peaks between two biological replicates of AUTS2 ChIP-seq in mouse brain.

a

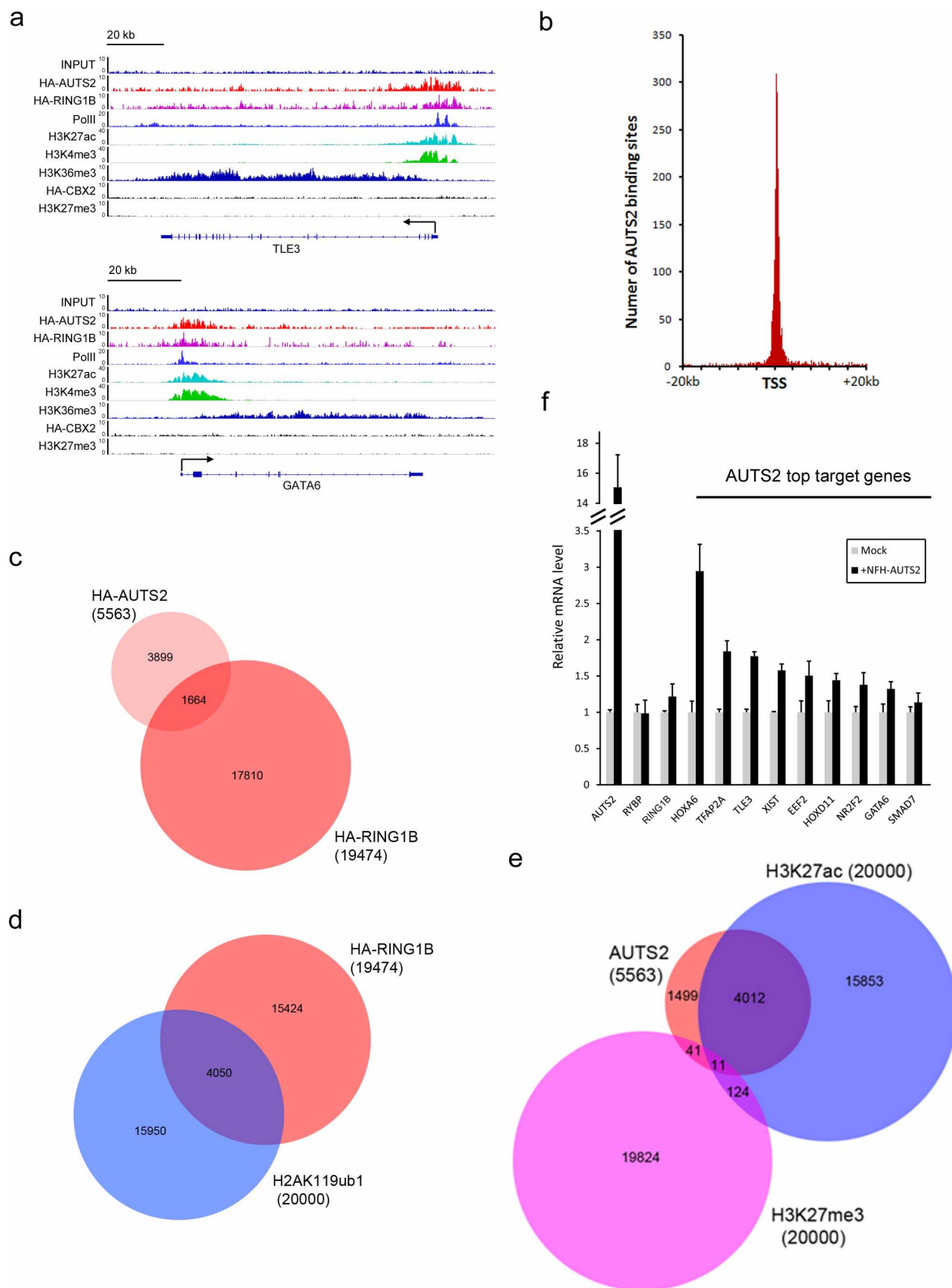


b



Extended Data Figure 7 | P300 is localized to AUTS2 targeted loci in mouse brain. **a**, IGV browser views for input, P300, AUTS2, Pol II, H3K27ac, H3K4me3, H3K36me3, H3K27me3, and H2AK119ub1 ChIP-seq performed in P1 mouse brain at two representative loci. The y axis corresponds to ChIP-seq

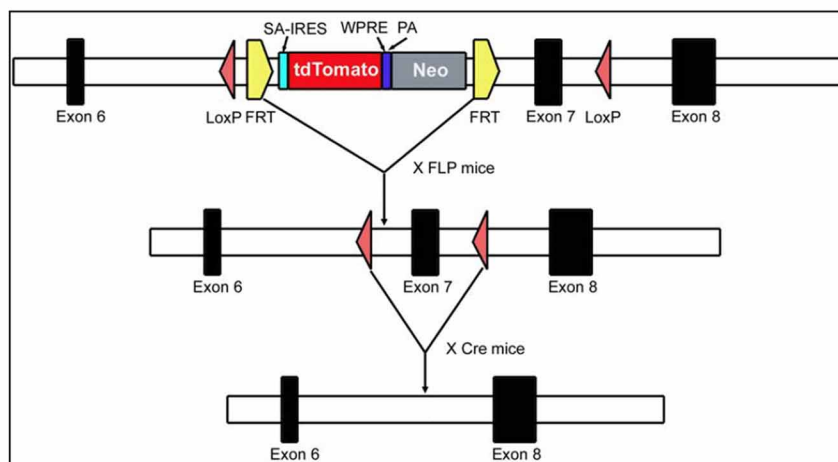
signal intensity. Gene representation at each locus is shown at the bottom. **b**, ChIP reads density plots for levels of P300 at loci co-targeted by AUTS2/RING1B and BMI1/RING1B. A ± 1 kb window relative to the centre of peaks is shown.



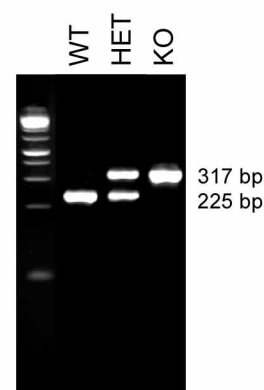
Extended Data Figure 8 | ChIP-seq in 293 T-REx cells. **a**, IGV browser views for input, HA-AUTS2, HA-RING1B, Pol II, H3K27ac, H3K4me3, H3K36me3, HA-CBX2, and H3K27me3 ChIP-seq libraries at two representative loci. The y axis corresponds to the ChIP-seq signal intensity. Gene representation at each locus is shown at the bottom. ChIP-seq data for HA-RING1B, HA-CBX2, and H3K27me3 obtained from a previous study¹⁶. **b**, Genomic distribution of HA-AUTS2 target regions relative to TSS. The x axis corresponds to the

distance from TSS (−20 kb to +20 kb); the y axis corresponds to frequency. **c–e**, Venn diagrams showing the overlap among regions targeted by factors as indicated. **f**, Analysis of mRNA levels of top targets identified by HA-AUTS2 ChIP-seq in 293 T-REx cells. RT-qPCR using the primers indicated was performed from vector control (mock) or NFH-AUTS2 stable cell lines induced by doxycycline (+NFH-AUTS2). All values are the mean of three technical replicates and error bars represent standard deviation.

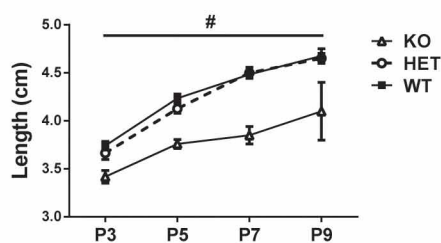
a



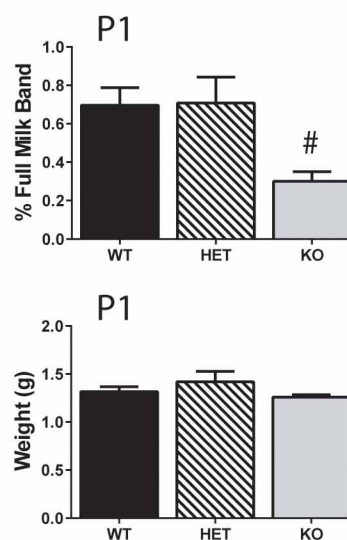
b



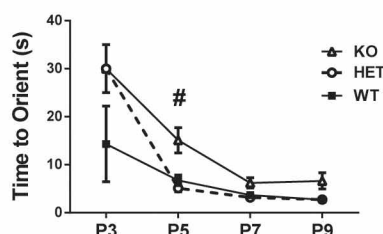
c



e



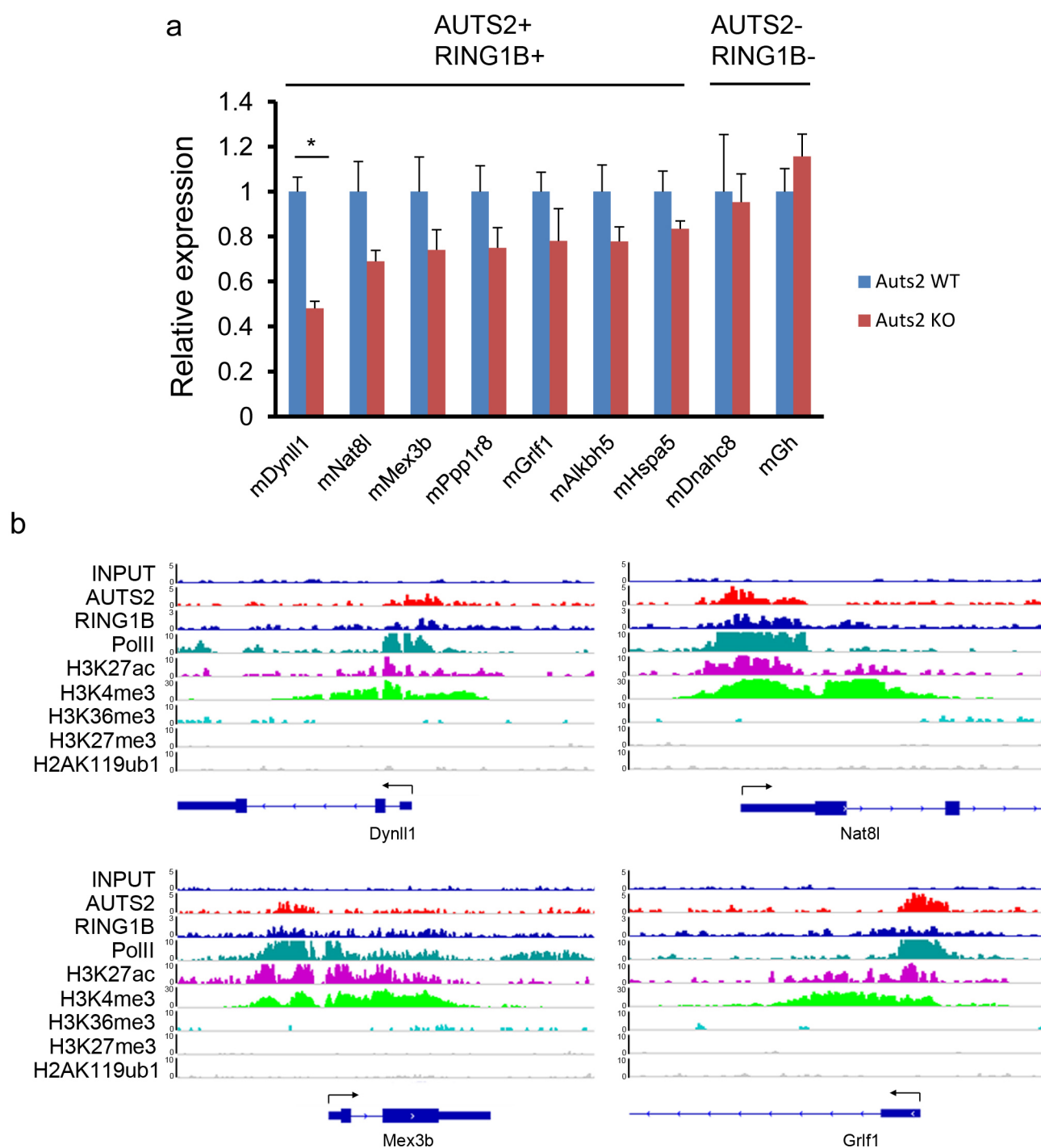
d



Extended Data Figure 9 | Generation of mice with *Auts2* conditional knockout in the nervous system and additional developmental phenotypes.

a, ES cells carrying an engineered allele of *Auts2* were generated through homologous recombination. Specifically, two *LoxP* sites were placed flanking exon 7 of *Auts2*. A cassette containing SA-IRES-tdTomato and an inverted PGK-Neo (neomycin phosphotransferase gene) were flanked by two *FRT* (FLP recombinase target) sites and inserted between the first *LoxP* site and exon 7. A WPRE (woodchuck hepatitis post-transcriptional regulatory element) sequence was placed immediately downstream of tdTomato to enhance its expression. Homologous mice carrying this engineered sequence are expected to give rise to a transcript containing only the first six exons of *Auts2* followed by IRES-driven tdTomato. Red fluorescence serves as a marker for

successful gene targeting. To obtain the conditional deletion of *Auts2*, these mice were crossed with FLP mice to excise the *FRT* flanking sequence, resulting in floxed mice, which were then crossed with nestin-Cre deleter mice to generate *Auts2* deletion in the nervous system. **b**, Genotyping of the *Auts2* floxed mice by PCR. The fast migrating species of 225 bp represents the PCR product of wild type, and the species of 317 bp corresponds to the knockout. **c**, Knockout mice are significantly shorter than both heterozygous and wild-type mice across development. #Post-hoc difference ($P < 0.05$) between wild type and knockout. **d**, The KO mice took significantly longer to orient their nose to an upward position. **e**, No significant difference in body weight was detected at P1, however, a significantly reduced milkband was observed in *Auts2* knockout.



Extended Data Figure 10 | Altered expression of genes targeted by PRC1-AUTS2 in brains of *Aut2* knockout mice. **a**, Expression profiles of select genes simultaneously targeted by AUTS2 and RING1B (labelled as AUTS2+ RING1B+). As negative control, two non-target genes were used (labelled as AUTS2- RING1B-). Total RNAs were extracted from whole brains of either wild-type or knockout mice, followed by reverse transcription

to generate cDNAs for RT-qPCR. Expression levels are normalized over those in wild type. All mean values of expression levels and standard errors were calculated from duplicated measurements of three biological replicates. * $P < 0.05$ by two-sided t -test. **b**, IGV views of four representative loci for genes examined as in **a**, showing the enrichment of AUTS2, RING1B, Pol II, and active histone marks.

Structure of influenza A polymerase bound to the viral RNA promoter

Alexander Pflug^{1,2*}, Delphine Guigay^{1,2*}, Stefan Reich^{1,2*} & Stephen Cusack^{1,2}

The influenza virus polymerase transcribes or replicates the segmented RNA genome (viral RNA) into viral messenger RNA or full-length copies. To initiate RNA synthesis, the polymerase binds to the conserved 3' and 5' extremities of the viral RNA. Here we present the crystal structure of the heterotrimeric bat influenza A polymerase, comprising subunits PA, PB1 and PB2, bound to its viral RNA promoter. PB1 contains a canonical RNA polymerase fold that is stabilized by large interfaces with PA and PB2. The PA endonuclease and the PB2 cap-binding domain, involved in transcription by cap-snatching, form protrusions facing each other across a solvent channel. The 5' extremity of the promoter folds into a compact hook that is bound in a pocket formed by PB1 and PA close to the polymerase active site. This structure lays the basis for an atomic-level mechanistic understanding of the many functions of influenza polymerase, and opens new opportunities for anti-influenza drug design.

Influenza A virus (FluA) mainly infects water and domestic fowl, although some strains cause disease in mammals such as humans, pigs, horses, seals and bats. The viral genome, composed of eight segments of negative-sense single-stranded RNA packaged in separate ribonucleoprotein particles, is transcribed and replicated by the heterotrimeric viral RNA (vRNA)-dependent RNA polymerase (RdRp), which comprises subunits PA, PB1 and PB2. The high mutation rate of the polymerase and the generation of novel viruses through reassortment of genome segments between different strains ensure rapid evolution of the virus with resultant seasonal epidemics and occasional, potentially devastating, pandemics. Although the polymerase has been studied extensively since the late 1960s, detailed understanding of its many functions both *in vitro* and in the context of the infected cell remains elusive (reviewed in refs 1 and 2), largely owing to the lack of atomic resolution structural information on the full-length polymerase. Nevertheless, in recent years, several crystal structures of fragments of the polymerase subunits have yielded important insights (reviewed in refs 1 and 3). These include the two domains involved in the unique cap-snatching mechanism of transcription used by the virus⁴—the PA amino-terminal endonuclease domain (PA-Nter)^{5,6}, and the central PB2 cap-binding domain⁷—structures that have contributed to a renaissance in anti-influenza drug design targeting the polymerase^{8,9}. In addition, structures are available of the inter-subunit interfaces between the PA carboxy-terminal domain (PA-C) and PB1-Nter (refs 10,11), between PB1-Cter and PB2-Nter (ref. 12), and of the PB2 C-terminal double 627-NLS domain¹³, which carry the host-specific PB2 residue 627 (Lys and Glu in human and avian strains, respectively) (reviewed in ref. 14) and the PB2 nuclear localization signal (NLS)¹⁵, respectively.

Here we describe the crystal structure of the complete heterotrimeric FluA polymerase bound to the vRNA promoter. To bypass difficulties in expression of recombinant human or avian polymerases, we used polymerase from the recently discovered bat-specific influenza virus (bat FluA)¹⁶, which is evolutionarily close to human/avian A strains with 70.0 (78.2), 79.5 (87.7) and 68.0 (78.6) per cent identity (similarity) for PA, PB1 and PB2, respectively (Supplementary Fig. 1). Bat polymerase can replicate efficiently in human cells¹⁶ and vice versa¹⁷, suggesting that the

bat structure will be a good model for all FluA polymerases. Here we describe the overall architecture of the polymerase, the structure of each subunit and their interfaces, and how the conserved 3' and 5' sequences of the vRNA promoter are bound. In the accompanying manuscript¹⁸, using two additional crystal structures of influenza B polymerase, implications of the structures for the mechanisms of *de novo* vRNA replication and cap-dependent transcription are presented.

Structure determination and overall architecture

Heterotrimeric influenza polymerase from A/little yellow-shouldered bat/Guatemala/060/2010(H17N10) was expressed in insect cells as a self-cleaving polyprotein and purified in milligram quantities to homogeneity (Extended Data Fig. 1). Using short templates, such as a 39-nucleotide vRNA mini-panhandle containing the conserved extremities or separated 3' (template) and 5' (activator) sequences, the recombinant bat polymerase is active in cap-dependent transcription as well as ApG-primed and, less efficiently, unprimed replication assays (Extended Data Fig. 2) without the need for the viral nucleoprotein, consistent with previous work¹⁹. Co-crystals of FluA polymerase were obtained with nucleotides 1–16 from the vRNA 5' end (5'-pAGUAGUAACAAGAGG G-3'), and nucleotides 1–18 or 3–18 from the 3' end (3'-OH-UCGU CUUCGUCUCCAUAU-5'-OH). The structure was solved by molecular replacement at 2.65 Å resolution using the structure of FluB polymerase¹⁸ (Extended Data Table 1). The FluA polymerase structure is 97.8% complete with 699 out of 714 (for PA), 750 out of 756 (for PB1), and 733 out of 760 (for PB2) residues modelled (2,182 out of 2,230 total).

The FluA polymerase has a U-shaped structure, with approximate height, width and depth of 115 × 100 × 75 Å, respectively (Fig. 1, Extended Data Fig. 3 and Supplementary Videos 1 and 2). The two protruding arms are formed by the PA-Nter endonuclease and PB2 cap-binding domains, which face each other across a solvent channel. The bottom of the U is formed by the large PA-C domain and one of the sides by the C-terminal two-thirds of PB2 (PB2-C) including the cap-binding domain. The body of the trimer is formed by PB1, decorated on one side by the N-terminal third of PB2 (PB2-N) (Fig. 1a, b) and on the other side by the linker (PA-linker) that connects the PA endonuclease (PA-Nter)

¹European Molecular Biology Laboratory, Grenoble Outstation, 71 Avenue des Martyrs, CS 90181, 38042 Grenoble Cedex 9, France. ²University Grenoble Alpes-Centre National de la Recherche Scientifique-EMBL Unit of Virus Host-Cell Interactions, 71 Avenue des Martyrs, CS 90181, 38042 Grenoble Cedex 9, France.

*These authors contributed equally to this work.

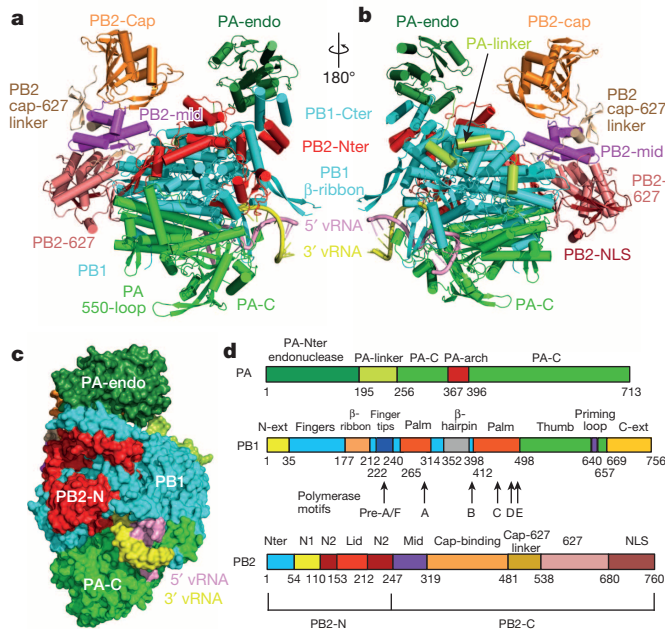


Figure 1 | Overall structure of the bat influenza A polymerase complex with the vRNA promoter. **a, b,** Two ribbon views colour-coded according to the domain structure in **d**, except that PA-C, PB1 and PB2-N are uniformly green, cyan and red, respectively. The vRNA 5' and 3' ends are pink and yellow tubes, respectively. **c,** Side-view in space-filling representation showing emergence of vRNA duplex at the interface of all three subunits. **d,** Subunit domain structure with subdomain names and colour scheme and showing the location of the conserved polymerase motifs in PB1.

with PA-C (Fig. 1b). Previous studies have revealed crucial but limited tail (Cter) to head (Nter) interactions between PA and PB1 (refs 10 and 11) and PB1 and PB2 (refs 12, 20 and 21). The actual inter-subunit interactions are much more extensive than this owing to an extremely complex intertwining of the subunits. The total buried surface area between PB1 and PA is 17,330 Å² and between PB1 and PB2 is around 14,100 Å², whereas the area between PA and PB2 is only 2,880 Å², confirming the central scaffolding role of PB1. The trimer contains a large, internal, catalytic and RNA-binding cavity formed by PB1 and PB2-N that is partially open at the top to the solvent channel between the PA endonuclease and PB2 cap-binding domains (putative template/product exit channel), as well as being accessible via two narrow side tunnels, the putative NTP and template entrance channels (see below). For sequence alignments of bat and human FluA polymerase and secondary structure assignments, see Supplementary Fig. 1. A schematic of each subunit domain structure is given in Fig. 1d.

PB1 subunit

Apart from the 15 N-terminal and 80 C-terminal residues, which form tight inter-subunit contacts with PA-C (refs 10, 11) and PB2-N (ref. 12), respectively, the detailed structure of the PB1 subunit has until now been completely unknown. However, sequence analysis revealed the presence of motifs pre-A (also known as F) and A–E characteristic of RNA-dependent RNA polymerases^{22–24} and correspondingly PB1 contains in its central region (residues 21–669) a typical right-handed RdRp fold, comprising fingers, fingertips, palm and thumb domains (Fig. 2a, b). A three-dimensional similarity search shows that hepatitis C virus (HCV) polymerase is structurally most like the polymerase region of PB1 (Fig. 2c), but many other RNA virus polymerases are also similar. Structural analysis has shown that *Flaviviridae* polymerases (for example, HCV, Dengue virus, West Nile virus)^{25–27} as well as bacteriophage Φ6 (ref. 28) contain a ‘priming loop’ to promote initiation of unprimed RNA synthesis²⁹. In PB1, residues 641–657 form a conserved anti-parallel β-loop (Fig. 2b) structurally analogous to the HCV priming loop (Fig. 2d), which could be involved in unprimed genome or anti-genome replication by influenza polymerase.

There are several idiosyncratic features of PB1. First, there are the N- and C-terminal extensions (N-ext and C-ext; Fig. 1d) that make inter-subunit contacts with PA and PB2, respectively. Second, there is an unusually long (~55 Å), solvent-exposed, flexibly hinged β-ribbon (strands β6 and β7, residues 177–212) (Fig. 2a, b). Interestingly, this element contains the PB1-NLS motifs, two separated basic patches (NSL1, 187-Lys/Arg-Lys-Lys/Arg-Arg-190 (bat/human) on β6; NSL2, 207-Lys-Lys-Arg/Lys-Val/Gln-Lys/Arg-211 on β7; Fig. 2a) that have been shown to be important for binding RanBP5, the PA–PB1 heterodimer nuclear import factor³⁰. A third special feature of PB1 is a β-hairpin insertion (strands β12 and β13, residues 352–360; Fig. 2a) in the finger domain, which, notably, is inserted through an extended loop in PA (the ‘PA-arch’; Fig. 3a). Both structures form an integral part of the 5' vRNA-binding site (see below). The C-terminal extension of PB1 after the putative priming loop is involved in direct 3'-template binding (residues 671–676, see below).

PA and PB2 subunits

The two structurally known domains of PA, the PA-Nter endonuclease domain (residues 1–195) and the large PA-C domain (258–714), are on opposite sides of the molecule, connected by the previously uncharacterized PA-linker (196–257) (Figs 1b and 3a), which wraps around the external face of the PB1 fingers and palm domain. In particular, residues 201–257, which include three helical segments (α7–α9), lie across the surface of PB1 making numerous, often conserved, inter-subunit contacts that are both hydrophobic and polar in nature (Extended Data Fig. 4a). The endonuclease domain is anchored to the rest of the polymerase through contacts with the same helical region of PB1-Cter that interacts with PB2-Nter, so that all three subunits are involved in

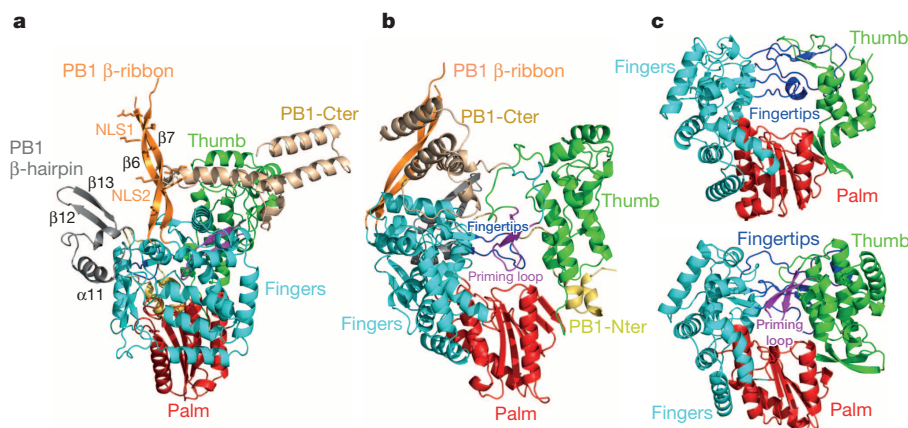


Figure 2 | PB1 structure and comparison with other RNA virus polymerases. **a,** Ribbon diagram of PB1, coloured as in Fig. 1d, highlighting idiosyncratic elements including the PB1-Cter extension (wheat), the β-ribbon (orange, with NLS1 and NLS2 motifs shown) and the β-hairpin (grey). **b,** As in **a** but rotated roughly 90° to show the internal cavity occupied by the putative priming loop (residues 640–657, magenta) and the PB1-Nter extension (yellow-orange). **c,** Same view as in **b** of Norwalk virus (PDB code 3BSO; top) and HCV (PDB code 2XI3; bottom) polymerases after superposition with PB1, and coloured equivalently. Norwalk and HCV polymerases both have two fingertip loops (blue) but only HCV has a priming loop.

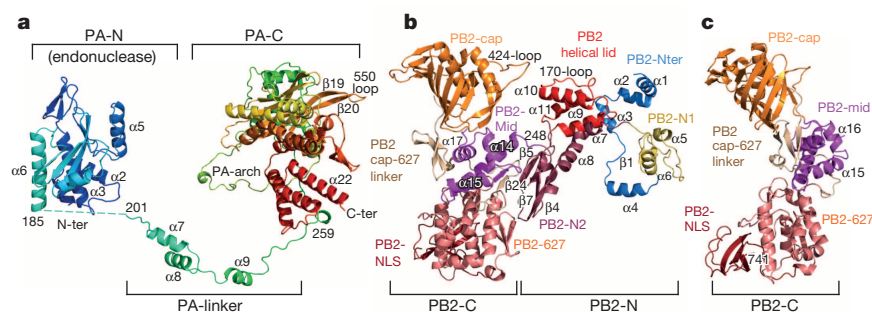


Figure 3 | PA and PB2 structure and the PA-linker–PB1 interface. **a**, The PA subunit in rainbow colouring from N-ter (dark blue) to C-ter (red). The PA-linker, PA-arch and 550-loop, which contains a putative host-specific residue, are highlighted. **b**, Ribbon diagram of the PB2 subunit with sub-domains coloured as in Fig. 1d. **c**, As in **b** but rotated roughly 90° and showing only the arc of the PB2-C domain.

positioning the endonuclease (Fig. 1a, b). The main contacts are via the packing of endonuclease helix $\alpha 4$ against both the penultimate PB1 helix $\alpha 21$ and the PB2 ‘170-loop’ (169–174), and via the endonuclease insertion (67–74) with the last PB1 helix $\alpha 22$ (Extended Data Fig. 4b). The endonuclease active site is solvent-exposed and facing the cap-binding domain (Fig. 1a, b), as discussed elsewhere in relation to the mechanism of cap-snatching¹⁸.

The PB2 subunit is divided into the N-terminal third (PB2-N, residues 1–247) and the C-terminal two-thirds (PB2-C, residues 248–760), each formed by several folded subdomains (Figs 1d and 3b, c). PB2-N comprises a series of linked modules that wrap around one edge and face of PB1, interacting mainly with the PB1 C-terminal extension and the polymerase thumb domain, opposite to where the PA linker binds (Figs 1 and 3b). After the well-characterized helical bundle interface with PB1-Cter, residues 35–54 of PB2-Nter are in an extended conformation followed by helix $\alpha 4$ that interacts with the template as it enters the polymerase active site (see below). Residues 55–103 ($\beta 1$, $\alpha 5$, $\beta 2$, $\beta 3$ and $\alpha 6$) form a more compact subdomain (PB2-N1) that buttresses the PB1 thumb domain (for example, PB2 helix $\alpha 6$ packs parallel against PB1 helix $\alpha 17$). Another linker leads to the PB2-N2 subdomain (residues 110–247), which has an extended shape (Fig. 3b). At one end a helical bundle ($\alpha 9$ – $\alpha 11$, residues 160–212) is inserted, denoted the PB2 helical lid. This includes the 170-loop (around 169–174), which contacts the endonuclease (Extended Data Fig. 4b), and the projecting helix $\alpha 10$, the N terminus (residue Asp 180) of which closely approaches the cap-binding domain. At the other extremity of the N2 domain are two anti-parallel β -ribbons ($\beta 4$ – $\beta 7$ and $\beta 5$ – $\beta 6$) with a helix inserted between them ($\alpha 12$ – $\alpha 13$). These make hydrophobic contacts with PA-Cter and with the thumb and palm domains of PB1.

PB2-C (residues 248–736) forms a single, arc-shaped unit (Fig. 3c), divided into five sub-domains, which constitutes one arm of the polymerase U-shape (Fig. 1). At one end of the arc is the cap-binding domain (319–481), and, at the other end, is the NLS domain (685–760), which is disordered beyond the NLS1 motif (736-Lys-Arg-Lys-Arg)¹⁵. The NLS domain is juxtaposed to the 627-domain (539–675) as observed in crystal structures of the isolated double 627-NLS domain^{13,31}. The loop carrying the host-specific residue 627, normally lysine in human and glutamate in avian strains but serine in bat, is in a solvent-exposed position remote from the PB1 active site. A possible role of the 627-domain is discussed elsewhere¹⁸ (see also Supplementary Information). The central part of the PB2-C arc is composed of two disconnected but interacting sub-domains: the PB2 mid-domain (248–319) that directly precedes the cap-binding domain, and the cap-627 linker (483–538). The mid-domain is a four helix bundle with one of the inter-helical linkers containing a short β -strand ($\beta 8$) that makes a stabilizing two-stranded parallel sheet with the cap-627 linker ($\beta 24$) (Fig. 3b). The bat cap-binding domain is very similar to that of human or avian Flua³², but Phe 357 forms one side of the methylated base sandwich rather than a histidine (Supplementary Fig. 1). The cap-627 linker proceeds from the C terminus of the cap-binding domain into a small three-stranded β -sheet (495–515, $\beta 21$ – $\beta 23$) that packs on the last helix ($\alpha 17$) of the PB2 mid-domain. This sheet has a distinctly concave, solvent-facing surface that could be involved in protein–protein interactions. The mid,

cap and cap-627 linker domains do not make extensive interfaces with other polymerase subunits.

PB1 functional regions

The catalytic centre responsible for template-directed nucleotide addition is located in the PB1 internal cavity and formed mainly by the highly conserved RdRp motifs pre-A/F and A–E. Comparison with known polymerase structures allows modelling of the template, substrate RNA and incoming NTPs into the PB1 active site, and deduction of the roles of certain key conserved residues (Fig. 4 and Extended Data Fig. 5). Motif pre-A/F is partly contained in the fingertips, a loop (residues 222–246) that extends from the fingers towards the thumb domain and the tip of which is stabilized by contacts with PA helix $\alpha 20$ (Fig. 2b and Extended Data Fig. 5a). Whereas HepC and Norwalk virus polymerases have two fingertip loops (one corresponding to motif F and the other closer to the polymerase N terminus) (Fig. 2c, d), influenza polymerase PB1-Nter is analogous to the second loop with residues 24–38 crossing from thumb to fingers in intimate association with the fingertips. Several conserved basic residues from motif pre-A/F are likely to be involved in template binding, and NTP channelling and binding³³ (Fig. 4a). Motif A contains the conserved active site Asp 305, which, together with Asp 445 and Asp 446 on motif C, coordinate two divalent metal ions (Fig. 4a) and promote catalysis³³. These residues have been shown to be essential for PB1 activity²³. Motif B has a characteristic methionine-rich loop in PB1 (406-GMMGMF), and is probably involved in stabilizing the base pair between the incoming NTP and the template. Motif D contains conserved Lys 480 and Lys 481 residues (involved in NTP binding) and is stabilized by contacts with PA helix $\alpha 20$ (656–663) and the PA peptide 671–684. Motif E forms another β -hairpin containing conserved residues thought to stabilize the position of the substrate/priming NTP (Fig. 4a).

As in other polymerases, a narrow tunnel, lined with positively charged residues, connects the internal cavity to the outside and this is presumed to attract and channel NTPs into the active site electrostatically (Extended Data Fig. 5a, b). In PB1, this putative NTP tunnel directly leads to the tip of the putative priming loop and involves highly conserved PB1 basic residues Arg 45, Lys 235, Lys 237 and Arg 239 (motif F3), Lys 308 (motif A), and Lys 480 and Lys 481 (motif D). A second tunnel constitutes the putative template entrance channel that is lined by conserved residues from all three subunits (Extended Data Fig. 5c, d).

Promoter binding

For initiation of RNA synthesis, the influenza polymerase needs to be bound to a promoter that comprises both conserved extremities of the pseudo-circularized vRNA or complementary RNA (cRNA)^{34,35}. The pyrimidine-rich 3′ (template) and purine-rich 5′ (activator) extremities are partially complementary and can form a non-canonical double helix, usually referred to as the panhandle³⁶. However, they are thought to bind the polymerase in a partially single-stranded conformation³⁵, either as a ‘corkscrew’^{37,38} or a ‘fork’^{39,40}, or as a combination of both⁴¹. These models concur on the presence of a distal base-paired region between nucleotides 11–14 of the 5′ and 10–13 of the 3′ ends, but differ in whether the individual proximal strands have internal structure or not.

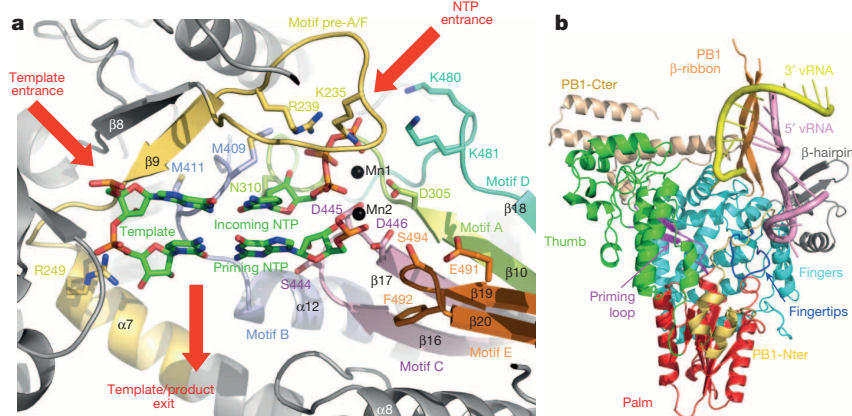


Figure 4 | PB1 functional regions. **a**, View into the PB1 catalytic site showing conserved polymerase motifs and key functional residues colour-coded according to: motif pre-A/F (residues 229–257, orange-yellow), motif A (296–314, lime), motif B (401–422, light blue), motif C (436–449, magenta), motif D (474–486, green-cyan) and motif E (487–497, orange). Template, substrate/priming nucleotide and incoming NTP (green) and two divalent cations (black spheres, coordinated by Asp 305, Asp 445 and Asp 446) are modelled after superposition with the Norwalk polymerase primer-template structure (PDB code 3BSO). Directions of NTP and template entrance tunnels and the template/product exit are indicated by arrows. Motif A contains the conserved

active site Asp 305 as well as Asn 310 (probably to bind the 2' OH of incoming NTP) and Lys 308 (NTP tunnel). Motif B has a characteristic methionine-rich loop (406-GMMMGMF) and probably stabilizes the base pair between the incoming NTP and the template. Motif C forms a β -hairpin containing Ser 444 (2' OH of priming NTP) and active site aspartates Asp 445 and Asp 446. Motif D contains conserved Lys 480 and Lys 481 residues (NTP channel). Motif E forms another β -hairpin containing conserved Glu 491, Phe 492 and Ser 494; it probably stabilizes the position of the substrate/priming nucleotide. **b**, Context of the vRNA promoter relative to the PB1 polymerase domain. PB1 is coloured as in Fig. 1d.

The polymerase–promoter crystal structure shows that the distal region is indeed base-paired, and that nucleotides 1–10 of the 5' end form a compact stem-loop (hook) structure (Fig. 4b).

The hook structure, formed by nucleotides 1–10 of the 5' vRNA (5'-pAGUAGUAACA), has two central canonical base pairs (G2–C9 and U3–A8) flanked by mismatch base pairs A1–A10 and A4–A7 (Fig. 5a). The stem is capped by G5, which is stacked antiparallel on A4 and U6 whose base faces outward. The sequence characteristics of the 5' hook are conserved in all known influenza virus vRNAs and cRNAs, the only variations, reflecting the imperfect complementarity of the two extremities, being the nature of the 2–9 and 3–8 Watson–Crick base pairs (G–C and A–U in vRNA, and G–C and C–G in cRNA, respectively) and the loop nucleotides 5 (usually a G) and 6 (usually an A). This hook structure is also likely to be conserved in orthomyxoviruses of the Thogoto lineage, except that G4–A7 would replace the A4–A7 mismatch⁴².

The 5' hook is sandwiched in a pocket formed on one side by strands β 17– β 18 and β 20 of the main β -sheet of PA, and on the other by the PA-arch (366–397) and the PB1 β -hairpin (353–370) that inserts through the arch (Fig. 5b). The buried surface area of the 5' end totals 4,044 Å² (60% with PA, 40% with PB1). Numerous polar interactions to the backbone (Extended Data Table 2) sense the shape of the stem-loop, including contacts to all phosphates (except 6–7) as well as to several ribose 2' OHs. Base contacts are made to invariant 5' residues G2, A7, A10 and A11 as well as to G5 and U6. Key interacting and highly conserved residues from PA are His 326, the peptide 366–370, 388-Tyr-Lys, 503-Arg-Leu-His, Lys 534, Arg 561 and Lys 569. From PB1 they include His 32, Thr 34 and Tyr 38 (conserved in all influenza strains) and 356-Met-Phe-Glu (Fig. 5c, d and Extended Data Fig. 6). An especially dense series of interactions binds and stabilizes the sharp turn between 5' A10–A11 (Fig. 5c). The PA-arch motif 366-Gly-Glu-Gln-Ala-370 forms a phosphate-binding loop, which interacts tightly with the backbone of A10–A11. His 505 (His 510 in human/avian strains) stacks on base A11 and hydrogen bonds to unpaired G9 of the 3' strand, which in turn stacks on PA Met 472. This histidine has previously been shown to be a crucial residue in regulating transcription⁴³. PA Arg 503 and PB1 Arg 365 make multivalent interactions with the RNA backbone (Fig. 5c). Conserved PB1 residues His 32 and Tyr 38 contact the phosphates of G5 and U6 and the double prolines 392-Pro-Pro in the PA-arch stack on the bases of these same nucleotides (Fig. 5d).

There are five base pairs in the duplex region of the promoter, 3' 10-UCUCC-14 with 5' 11-AGAGG-15, which projects away from the polymerase (Fig. 1c). The self-complementary four-nucleotide overhang 15-AUAU-18 of the crystallized 3' end base-pairs with a crystal symmetry-related equivalent, thus forming a pseudo-continuous double-stranded RNA of 14 base pairs between two two-fold-related polymerases (Extended Data Fig. 7). The duplex region of the promoter is contacted by the central section of the long PB1 β -ribbon and by residues 672–676 of PB1-Cter (Extended Data Fig. 6). The PA peptide 503-Arg-Leu-His, reinforced by 466–475, forms a wedge that separates the 5' and 3' strands into binding pockets (Extended Data Fig. 6). Only the proximal single-stranded 3' nucleotides 6-UUCG-9 are visible in the structure, and these are directed towards the polymerase template entry tunnel before turning away towards the solvent. There is a sharp turn between unpaired 3' end nucleotides G9 and C8 (Extended Data Fig. 6). Residues, very highly conserved in all influenza strains, from all three subunits (PA 505–509 and Lys 567, PB1-Cter 671–676 and PB2 36–49) are involved in binding the 3' nucleotides 6-UUCG-9 (Extended Data Fig. 6). At the apex of the sharp turn, the phosphate of 3' C8 is bound by PA Lys 567 and PB2 Arg 46, the latter being positioned by salt bridges with PA Asp 509 and PB2 Glu 40. PA Arg 507 and PB1 C-terminal extension residues Asn 671, Arg 672 and Ser 673 interact extensively with the backbone of 3' U7 and U10.

Conclusions

The structure of influenza polymerase, the first from any negative-strand RNA virus, reveals the enormous complexity of the molecule and highlights the fact that all three subunits are intricately involved in many of most important functional regions. This undoubtedly explains why 40 years' of polymerase biochemistry has often led to confusing and contradictory results. For instance, numerous studies have tried to identify the vRNA 3'- and 5'-end binding sites by crosslinking and/or mutagenesis^{44–47} but have failed to reveal the critical residues (see Supplementary Information). Conversely, the vRNA promoter structure itself is essentially as predicted⁴¹, although the A–A mismatches in the 5'-end hook were not foreseen. Indeed, the hook, tightly bound in a pocket formed by PA and PB1, is an integral part of the polymerase structure and this binding is required to enhance or activate polymerase functions^{48–50} (Extended Data Fig. 2). Without an apo-structure, this cannot

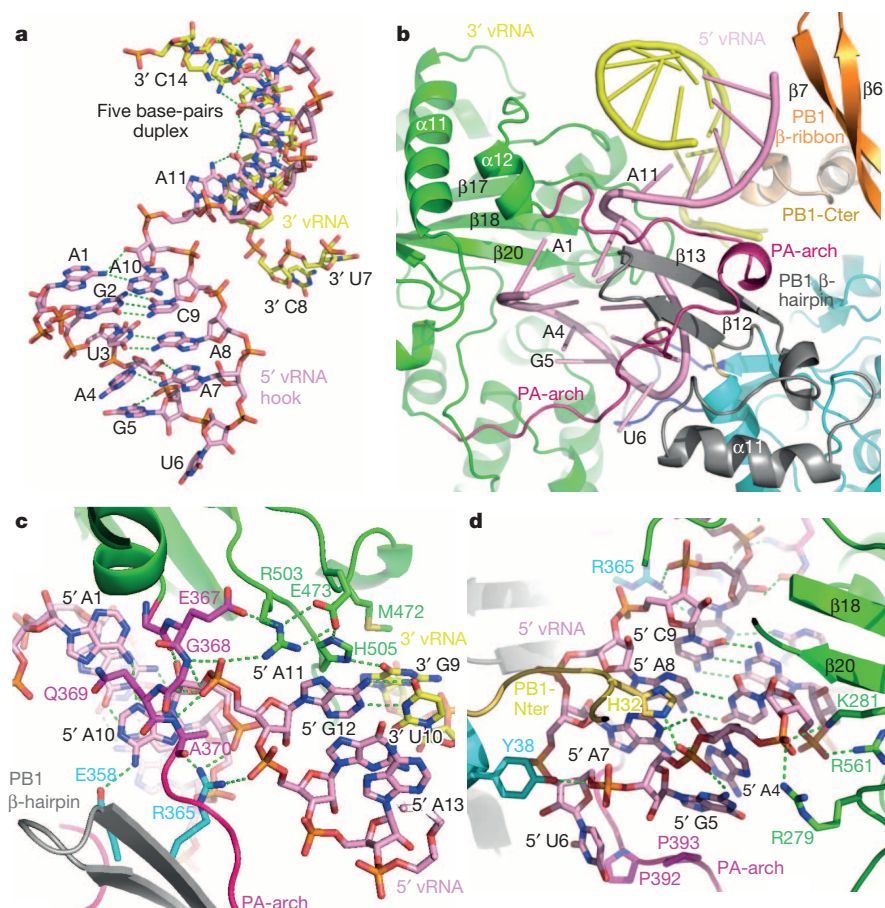


Figure 5 | Structure of the vRNA promoter and how it binds to the polymerase. **a**, Stick representation of the vRNA promoter highlighting internal hydrogen bonds (green dotted lines) within the 5' hook structure (pink) and the distal duplex region with the 3' end (yellow). The non-canonical A1–A10 and A4–A7 pairs are both of the N6 amino (A1, A7)–N3 (A4, A10) type. **b**, Ribbon diagram of the 5'-hook binding site between the PA β -sheet and PA-arch (plum) and the inserted PB1 β -hairpin (grey). PA is otherwise green and PB1 cyan. The 3'–5' duplex region contacts the PB1 β -ribbon (orange) notably via residues Lys 188, Thr 201 and Arg 203. His 505 stacks on base A11 of the 5' strand, and contacts the O6 of unpaired G9 of the 3' strand, which in turn stacks on PA Met 472. PA Arg 503 and the phosphate-binding loop (367–370) within the PA-arch (plum) interact with the phosphate of 5' A11. PB1 β -hairpin residue Arg 365 (cyan) makes hydrogen bonds to the phosphates of 5' nucleotides C9, A10 and G12 as well as to the N7 of A10, and Glu 358 (cyan) contacts the N6 of A10. **d**, Protein interactions of the 5' hook involving highly conserved PB1 N-terminal residues His 32 and Tyr 38, and PA basic residues Lys 281, Arg 279 and Arg 561. Pro 392 and Pro 393 from the PA-arch (plum) stack with 5' nucleotides U6 and G5, respectively; only the second proline is universally conserved in all influenza strains.

be fully rationalised yet, but it is likely that without the stabilization promoted by 5'-end binding the nearby polymerase active site will be disorganized. Whereas, in the bat polymerase structure, the 3' end of the template is not completely visible, in the FluB polymerase structure the complete 3' strand is well ordered¹⁸. However, rather than being directed in to the PB1 active site, the vRNA 3' end seems to have an alternative, but specific, binding site lying on the surface of the polymerase in the vicinity of the long PB1 β -ribbon. This is discussed further in the accompanying paper, along with other insights into polymerase function derived from the structure¹⁸.

There is considerable interest in understanding the exact role of polymerase residues that have been implicated in host adaptation, notably between avian and human influenza A strains¹⁴. Such mutations, identified by analysis of natural sequences or serial adaptation of viruses to mice, typically have a neutral effect in avian cells but enhance polymerase activity in mammalian cells. Because the positions of implicated residues can henceforth be mapped onto the full polymerase structure, an initial distinction can now be made between those residues that are more likely, because of their internal location, to affect the intrinsic rate of polymerase functions (which could be important for species-dependent physiological reasons), and others, which, because of their surface location, possibly act through direct interaction with other viral or cellular factors. Some initial observations are made in the Supplementary Information, but further structural studies of the polymerase in different functional conformations and eventually with bound host factors are required to determine the exact role of these putative host-specific residues.

Finally, the unexpectedly good resolution of this crystal structure gives hope that structure-based drug design targeting the PB1 active site, vRNA binding or numerous potential allosteric sites, will soon become possible.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 18 August; accepted 29 October 2014.

Published online 19 November 2014.

1. Resa-Infante, P., Jorba, N., Coloma, R. & Ortin, J. The influenza virus RNA synthesis machine: advances in its structure and function. *RNA Biol.* **8**, 207–215 (2011).
2. Fodor, E. The RNA polymerase of influenza A virus: mechanisms of viral transcription and replication. *Acta Virol.* **57**, 113–122 (2013).
3. Ruigrok, R. W., Crepin, T., Hart, D. J. & Cusack, S. Towards an atomic resolution understanding of the influenza virus replication machinery. *Curr. Opin. Struct. Biol.* **20**, 104–113 (2010).
4. Plotch, S. J., Bouloy, M., Ulmanen, I. & Krug, R. M. A unique cap(m⁷GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell* **23**, 847–858 (1981).
5. Dias, A. *et al.* The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **458**, 914–918 (2009).
6. Yuan, P. *et al.* Crystal structure of an avian influenza polymerase PA_N reveals an endonuclease active site. *Nature* **458**, 909–913 (2009).
7. Guilligay, D. *et al.* The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nature Struct. Mol. Biol.* **15**, 500–506 (2008).
8. Kowalinski, E. *et al.* Structural analysis of specific metal chelating inhibitor binding to the endonuclease domain of influenza pH1N1 (2009) polymerase. *PLoS Pathog.* **8**, e1002831 (2012).
9. Clark, M. P. *et al.* Discovery of a novel, first-in-class, orally bioavailable azaindole inhibitor (VX-787) of influenza PB2. *J. Med. Chem.* **57**, 6668–6678 (2014).
10. He, X. *et al.* Crystal structure of the polymerase PA_C–PB1_N complex from an avian influenza H5N1 virus. *Nature* **454**, 1123–1126 (2008).
11. Obayashi, E. *et al.* The structural basis for an essential subunit interaction in influenza virus RNA polymerase. *Nature* **454**, 1127–1131 (2008).
12. Sugiyama, K. *et al.* Structural insight into the essential PB1–PB2 subunit contact of the influenza virus RNA polymerase. *EMBO J.* **28**, 1803–1811 (2009).
13. Tarendeau, F. *et al.* Host determinant residue lysine 627 lies on the surface of a discrete, folded domain of influenza virus polymerase PB2 subunit. *PLoS Pathog.* **4**, e1000136 (2008).
14. Cauldwell, A. V., Long, J. S., Moncorge, O. & Barclay, W. S. Viral determinants of influenza A virus host range. *J. Gen. Virol.* **95**, 1193–1210 (2014).

15. Tarendeau, F. *et al.* Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. *Nature Struct. Mol. Biol.* **14**, 229–233 (2007).
16. Tong, S. *et al.* A distinct lineage of influenza A virus from bats. *Proc. Natl Acad. Sci. USA* **109**, 4269–4274 (2012).
17. Poole, D. S. *et al.* Influenza A virus polymerase is a site for adaptive changes during experimental evolution in bat cells. *J. Virol.* **88**, 12572–12585 (2014).
18. Reich, S. *et al.* Structural insights into cap-snatching and RNA synthesis by influenza virus polymerase. *Nature* <http://dx.doi.org/10.1038/nature14009> (this issue).
19. Turrell, L., Lyall, J. W., Tiley, L. S., Fodor, E. & Vreede, F. T. The role and assembly mechanism of nucleoprotein in influenza A virus ribonucleoprotein complexes. *Nature Commun.* **4**, 1591 (2013).
20. González, S., Zurcher, T. & Ortin, J. Identification of two separate domains in the influenza virus PB1 protein involved in the interaction with the PB2 and PA subunits: a model for the viral RNA polymerase structure. *Nucleic Acids Res.* **24**, 4456–4463 (1996).
21. Poole, E. L., Medcalf, L., Elton, D. & Digard, P. Evidence that the C-terminal PB2-binding region of the influenza A virus PB1 protein is a discrete α -helical domain. *FEBS Lett.* **581**, 5300–5306 (2007).
22. Müller, R., Poch, O., Delarue, M., Bishop, D. H. & Bouloy, M. Rift Valley fever virus L segment: correction of the sequence and possible functional role of newly identified regions conserved in RNA-dependent polymerases. *J. Gen. Virol.* **75**, 1345–1352 (1994).
23. Biswas, S. K. & Nayak, D. P. Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1. *J. Virol.* **68**, 1819–1826 (1994).
24. Bruenn, J. A. A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. *Nucleic Acids Res.* **31**, 1821–1829 (2003).
25. Yap, T. L. *et al.* Crystal structure of the dengue virus RNA-dependent RNA polymerase catalytic domain at 1.85-angstrom resolution. *J. Virol.* **81**, 4753–4765 (2007).
26. Lesburg, C. A. *et al.* Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nature Struct. Biol.* **6**, 937–943 (1999).
27. Bressanelli, S. *et al.* Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proc. Natl Acad. Sci. USA* **96**, 13034–13039 (1999).
28. Butcher, S. J., Grimes, J. M., Makeyev, E. V., Bamford, D. H. & Stuart, D. I. A mechanism for initiating RNA-dependent RNA polymerization. *Nature* **410**, 235–240 (2001).
29. Caillet-Saguy, C., Lim, S. P., Shi, P. Y., Lescar, J. & Bressanelli, S. Polymerases of hepatitis C viruses and flaviviruses: structural and mechanistic insights and drug development. *Antiviral Res.* **105**, 8–16 (2014).
30. Hutchinson, E. C., Orr, O. E., Man Liu, S., Engelhardt, O. G. & Fodor, E. Characterization of the interaction between the influenza A virus polymerase subunit PB1 and the host nuclear import factor Ran-binding protein 5. *J. Gen. Virol.* **92**, 1859–1869 (2011).
31. Kuzuhara, T. *et al.* Structural basis of the influenza A virus RNA polymerase PB2 RNA-binding domain containing the pathogenicity-determinant lysine 627 residue. *J. Biol. Chem.* **284**, 6855–6860 (2009).
32. Pautus, S. *et al.* New 7-methylguanine derivatives targeting the influenza polymerase PB2 cap-binding domain. *J. Med. Chem.* **56**, 8915–8930 (2013).
33. Zamyatkin, D. F. *et al.* Structural insights into mechanisms of catalysis and inhibition in Norwalk virus polymerase. *J. Biol. Chem.* **283**, 7705–7712 (2008).
34. Fodor, E., Pritlove, D. C. & Brownlee, G. G. The influenza virus panhandle is involved in the initiation of transcription. *J. Virol.* **68**, 4092–4096 (1994).
35. Tiley, L. S., Hagen, M., Matthews, J. T. & Krystal, M. Sequence-specific binding of the influenza virus RNA polymerase to sequences located at the 5' ends of the viral RNAs. *J. Virol.* **68**, 5108–5116 (1994).
36. Hsu, M. T., Parvin, J. D., Gupta, S., Krystal, M. & Palese, P. Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. *Proc. Natl Acad. Sci. USA* **84**, 8140–8144 (1987).
37. Neumann, G. & Hobom, G. Mutational analysis of influenza virus promoter elements *in vivo*. *J. Gen. Virol.* **76**, 1709–1717 (1995).
38. Flick, R., Neumann, G., Hoffmann, E., Neumeier, E. & Hobom, G. Promoter elements in the influenza vRNA terminal structure. *RNA* **2**, 1046–1057 (1996).
39. Fodor, E., Pritlove, D. C. & Brownlee, G. G. Characterization of the RNA-fork model of virion RNA in the initiation of transcription in influenza A virus. *J. Virol.* **69**, 4012–4019 (1995).
40. Kim, H. J., Fodor, E., Brownlee, G. G. & Seong, B. L. Mutational analysis of the RNA-fork model of the influenza A virus vRNA promoter *in vivo*. *J. Gen. Virol.* **78**, 353–357 (1997).
41. Pritlove, D. C., Poon, L. L., Devenish, L. J., Leahy, M. B. & Brownlee, G. G. A hairpin loop at the 5' end of influenza A virus virion RNA is required for synthesis of poly(A)⁺ mRNA *in vitro*. *J. Virol.* **73**, 2109–2114 (1999).
42. Briesse, T. *et al.* *Upolu virus* and *Aransas Bay virus*, two presumptive bunyaviruses, are novel members of the family *Orthomyxoviridae*. *J. Virol.* **88**, 5298–5309 (2014).
43. Fodor, E. *et al.* A single amino acid mutation in the PA subunit of the influenza virus RNA polymerase inhibits endonucleolytic cleavage of capped RNAs. *J. Virol.* **76**, 8989–9001 (2002).
44. González, S. & Ortin, J. Characterization of influenza virus PB1 protein binding to viral RNA: two separate regions of the protein contribute to the interaction domain. *J. Virol.* **73**, 631–637 (1999).
45. Li, M. L., Ramirez, B. C. & Krug, R. M. RNA-dependent activation of primer RNA production by influenza virus polymerase: different regions of the same protein subunit constitute the two required RNA-binding sites. *EMBO J.* **17**, 5844–5852 (1998).
46. Jung, T. E. & Brownlee, G. G. A new promoter-binding site in the PB1 subunit of the influenza A virus polymerase. *J. Gen. Virol.* **87**, 679–688 (2006).
47. Kerry, P. S., Willsher, N. & Fodor, E. A cluster of conserved basic amino acids near the C-terminus of the PB1 subunit of the influenza virus RNA polymerase is involved in the regulation of viral transcription. *Virology* **373**, 202–210 (2008).
48. Leahy, M. B., Pritlove, D. C., Poon, L. L. & Brownlee, G. G. Mutagenic analysis of the 5' arm of the influenza A virus virion RNA promoter defines the sequence requirements for endonuclease activity. *J. Virol.* **75**, 134–142 (2001).
49. Rao, P., Yuan, W. & Krug, R. M. Crucial role of CA cleavage sites in the cap-snatching mechanism for initiating viral mRNA synthesis. *EMBO J.* **22**, 1188–1198 (2003).
50. Poon, L. L., Pritlove, D. C., Sharps, J. & Brownlee, G. G. The RNA polymerase of influenza virus, bound to the 5' end of virion RNA, acts *in cis* to polyadenylate mRNA. *J. Virol.* **72**, 8214–8219 (1998).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank members of the ESRF-EMBL Joint Structural Biology Group for access to European Synchrotron Radiation Facility (ESRF) beamlines, staff of the European Molecular Biology Laboratory (EMBL) eukaryotic expression and high-throughput crystallization facilities within the Partnership for Structural Biology (PSB), D. Hart for help with construct design, and H. Malet for electron microscopy. This work was supported by ERC Advanced Grant V-RNA (322586) to S.C.

Author Contributions A.P. did protein expression, purification and crystallization, with help from S.R. and D.G. who also did activity assays. A.P. did X-ray data collection and, together with S.C., did crystallographic analysis. S.C. supervised the project and wrote the paper with input from the other authors.

Author Information Structure factors and coordinates for Bat FluA have been deposited in the Protein Data Bank (PDB) under accession 4WSB. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.C. (cusack@embl.fr).

METHODS

Construct. The influenza A/little yellow-shouldered bat/Guatemala/060/2010 (H17N10) polymerase heterotrimer was expressed as a self-cleaving polypeptide (Extended Data Fig. 1a). A codon-optimized synthetic construct (DNA2.0) with the composition $\text{GNH}_{\text{BstEII}} \text{GSGENLYFQ}_{\text{TEV}} \text{GSHHHHHHHH}_{8 \times \text{His-tag}} \text{GSGS-PA}$ (GenBank ID AFC35437.1) $\text{GSGENLYFQ}_{\text{TEV}} \text{GSGSGSGSG-PB1}$ (GenBank ID AFC35436) $\text{GSGENLYFQ}_{\text{TEV}} \text{GSGSGSGSG-PB2}$ (GenBank ID AFC35435.1) $\text{GWSHPQFEK}_{\text{Strep-tag}} \text{GGGSGGGSGSAWSHPQFEK}_{\text{Strep-tag}} \text{GRSG}_{\text{RsrII}}$ was cloned via BstEII and RsrII sites into the vector pKL-PBac⁵¹, which also contains coding sequences for tobacco etch virus (TEV) protease (5') and cyan fluorescent protein (CFP) (3'). (The TEV-site, His-tag and Strep-tag are underlined.)

Expression and purification. The bat FluA polymerase was produced in HighFive insect cells using the baculovirus expression system. Cells were collected by centrifugation, re-suspended in buffer A (50 mM Tris-HCl, 500 mM NaCl, 10% (v/v) glycerol and 5 mM β -mercaptoethanol, pH 8) supplemented with protease inhibitors (Roche, complete mini, EDTA-free), and lysed by sonication. Cell debris was spun off (30 min, 4 °C, 35,000g) and ammonium sulphate added to the clarified supernatant (0.5 g ml⁻¹) to force the protein out of solution. The precipitated protein was collected by centrifugation (30 min, 4 °C, 70,000g) and re-suspended in buffer A. After a final centrifugation step (30 min, 4 °C, 70,000g) the polymerase was purified from the fraction of soluble proteins via immobilized metal ion affinity chromatography and a strep-tactin resin (IBA, Superflow), using buffer A as running buffer in both cases. Fractions containing the target protein were pooled and diluted with an equal volume of buffer B (50 mM HEPES/NaOH, 10% (v/v) glycerol and 2 mM TCEP, pH 7.5) before loading on a heparin column (HiPrep Heparin HP, GE Healthcare). Polymerase was eluted by a gradient of buffer B supplemented with 1 M NaCl, concentrated, and subjected to size-exclusion chromatography (S200, GE Healthcare) in buffer C (50 mM HEPES/NaOH, 500 mM NaCl, 5% (v/v) glycerol and 2 mM TCEP, pH 7.5). Monomeric and RNA-free polymerase was concentrated, flash-frozen and stored at -80 °C. The typical yield of pure heterotrimer is about 1 mg l⁻¹ of insect cells.

Crystallization, data collection and structure solution. Polymerase protein in buffer C was adjusted to a concentration of 10 mg ml⁻¹, mixed in a 1:1 ratio with vRNA, which was an equimolar mixture of nucleotides 1–16 from the 5' end (5'-pAGUAGUAAACAAGAGGG-3') and nucleotides 1–18 or 3–18 from the 3' end (3'-OH-UCGUCUUCGUCUCAUUAU-5' OH) (IBA). Crystallization trials were performed by vapour diffusion at 4 °C using a Cartesian robot. The best crystals grew in mother liquor containing 0.7–1.5 M sodium/potassium phosphate at pH 5.0. For data collection, crystals were flash-frozen in well solution supplemented with 25% glycerol. Diffraction data were collected at 100 °K with an X-ray wavelength of 0.9763 Å on beamline ID23-1 of the European Synchrotron Radiation Facility equipped with a Pilatus 6M-F detector and integrated and scaled with XDS⁵². Initial phases were obtained by molecular replacement with the structure of the influenza B polymerase¹⁸. The model was improved by making use of the five known high-resolution structures of FluA polymerase fragments (endonuclease⁵³, PA-Cter-PB1-Nter (PDB codes 2ZN1 and 3CM8), PB1-Cter/PB2-Nter (PDB code 3A1G), PB2-cap

and 627-NLS domains (PDB code 2VY6). Refinement was performed with Refmac⁵⁴. A putative zinc ion is found bound between PB1 His 562 and PA Asp 421. Figures were drawn with Pymol⁵⁵. The vRNA and most protein regions have very good electron density apart from a few connecting peptides and the PA endonuclease domain, which has poor density except where it contacts the rest of the polymerase. Ramachandran statistics, as calculated by Molprobity⁵⁶ are 94.2% (favoured), 0.7% (disallowed).

Polymerase activity assays. A T7-transcribed 39-nucleotide mini-panhandle or equimolar mixture of separated synthetic 3' and 5' ends were used as vRNA (Extended Data Fig. 2a, b).

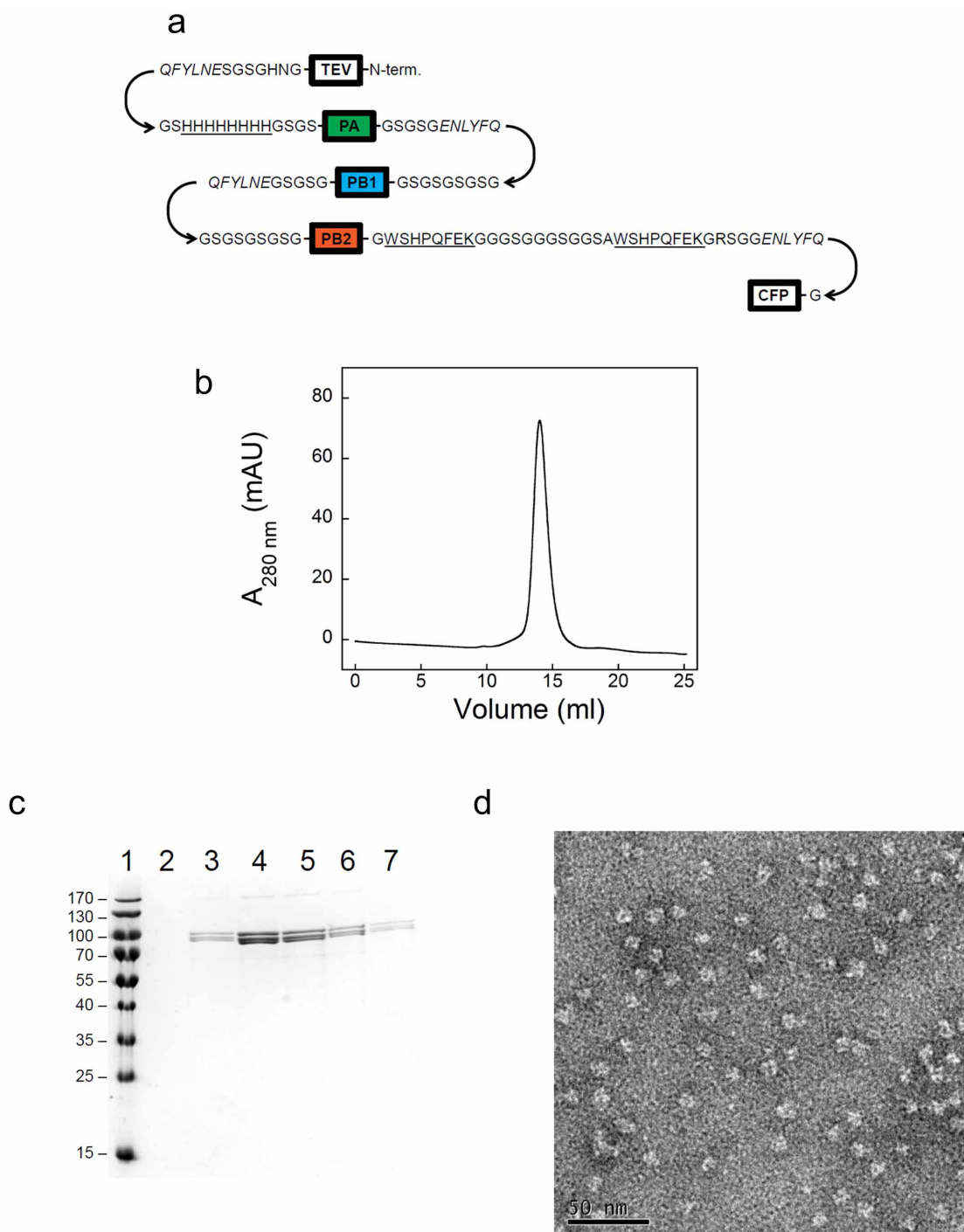
For the ApG-primed replication assay, 0.5 μM protein, 0.5 μM vRNA, 0.5 mM ApG, 0.4 mM GTP/CTP, 1 mM ATP, 0.04 mM UTP, ³²P-UTP and 0.8 U μl^{-1} Ribolock, in buffer (150 mM NaCl, 50 mM HEPES, pH 7.5, 5 mM MgCl₂ and 2 mM TCEP) were mixed and incubated at 30 °C for 2 h.

For the cap-dependent transcription assay, 0.5 μM protein, 0.5 μM vRNA, 0.4 mM GTP/CTP/UTP, 1 mM ATP and ³²P-labelled capped RNA in the same buffer (150 mM NaCl, 50 mM HEPES, pH 7.5, 5 mM MgCl₂ and 2 mM TCEP) were mixed and incubated at 30 °C for 2 h. For this purpose, a 5' diphosphate synthetic 20-base RNA, 5'-ppAAUCUAUAAUAGCAUUAUCC-3' (Chemgenes), was capped by incubating with vaccinia virus capping enzyme (purified in house following ref. 57) and 20 μM SAM, ³²P-GTP, 50 mM Tris, pH 8.0, 6 mM KCl, 1.25 mM MgCl₂ and 0.8 U μl^{-1} Ribolock.

For the endonuclease assay, transcription mix without any NTPs was incubated at 30 °C for 2 h. Samples were separated on 7 M urea, 20% acrylamide gel in TBE buffer, exposed on a storage phosphor screen and read with a Typhoon scanner.

For the time course of unprimed and ApG-primed vRNA replication, 0.5 μM bat FluA polymerase was mixed with 1 μM 39-nucleotide vRNA mini-panhandle template, NTPs (1 mM ATP, 0.4 mM GTP, 0.4 mM CTP and 0.04 mM UTP) and 0.12 $\mu\text{Ci} \mu\text{l}^{-1}$ ³²P-UTP, in the absence or presence of 0.5 mM ApG. Reactions were incubated at 30 °C and samples were analysed on a 20% acrylamide, 7 M urea denaturing gel after 0, 2, 5, 10, 15, 20, 30, 40 and 50 min, 1, 2 and 3 h.

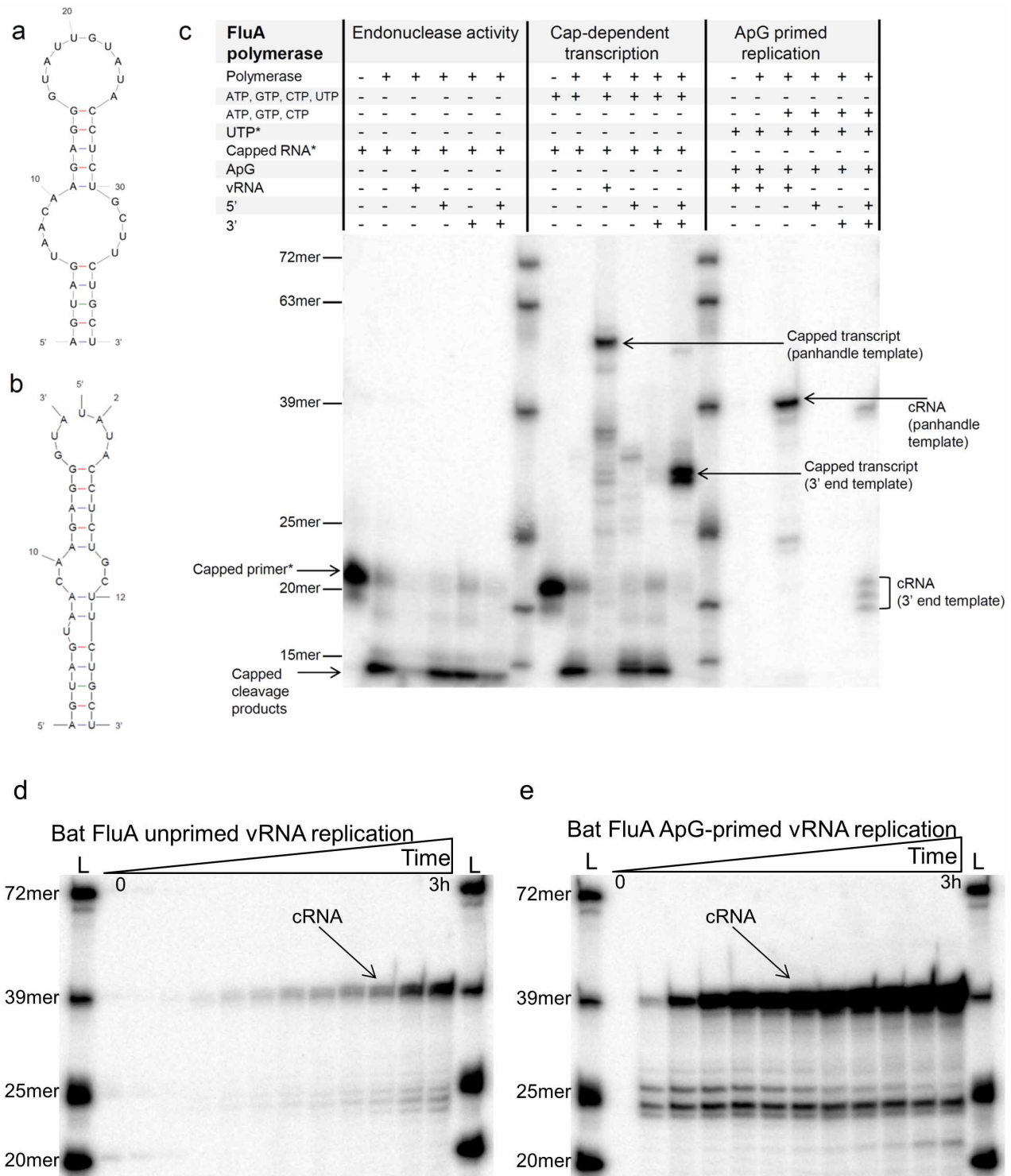
51. Nie, Y., Bellon-Echeverria, I., Trowitzsch, S., Bieniossek, C. & Berger, I. Multiprotein complex production in insect cells by using polyproteins. *Methods Mol. Biol.* **1091**, 131–141 (2014).
52. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D* **66**, 133–144 (2010).
53. Tefsen, B. *et al.* The N-terminal domain of PA from bat-derived influenza-like virus H17N10 has endonuclease activity. *J. Virol.* **88**, 1935–1941 (2014).
54. Murshudov, G. N. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
55. DeLano, W. L. The PyMOL Molecular Graphics System; <http://www.pymol.sourceforge.net> (Schrödinger, LLC, 2002).
56. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
57. De la Peña, M., Kyrieleis, O. J. & Cusack, S. Structural insights into the mechanism and evolution of the vaccinia virus mRNA cap N7 methyl-transferase. *EMBO J.* **26**, 4913–4925 (2007).



Extended Data Figure 1 | Production of influenza A polymerase heterotrimer.

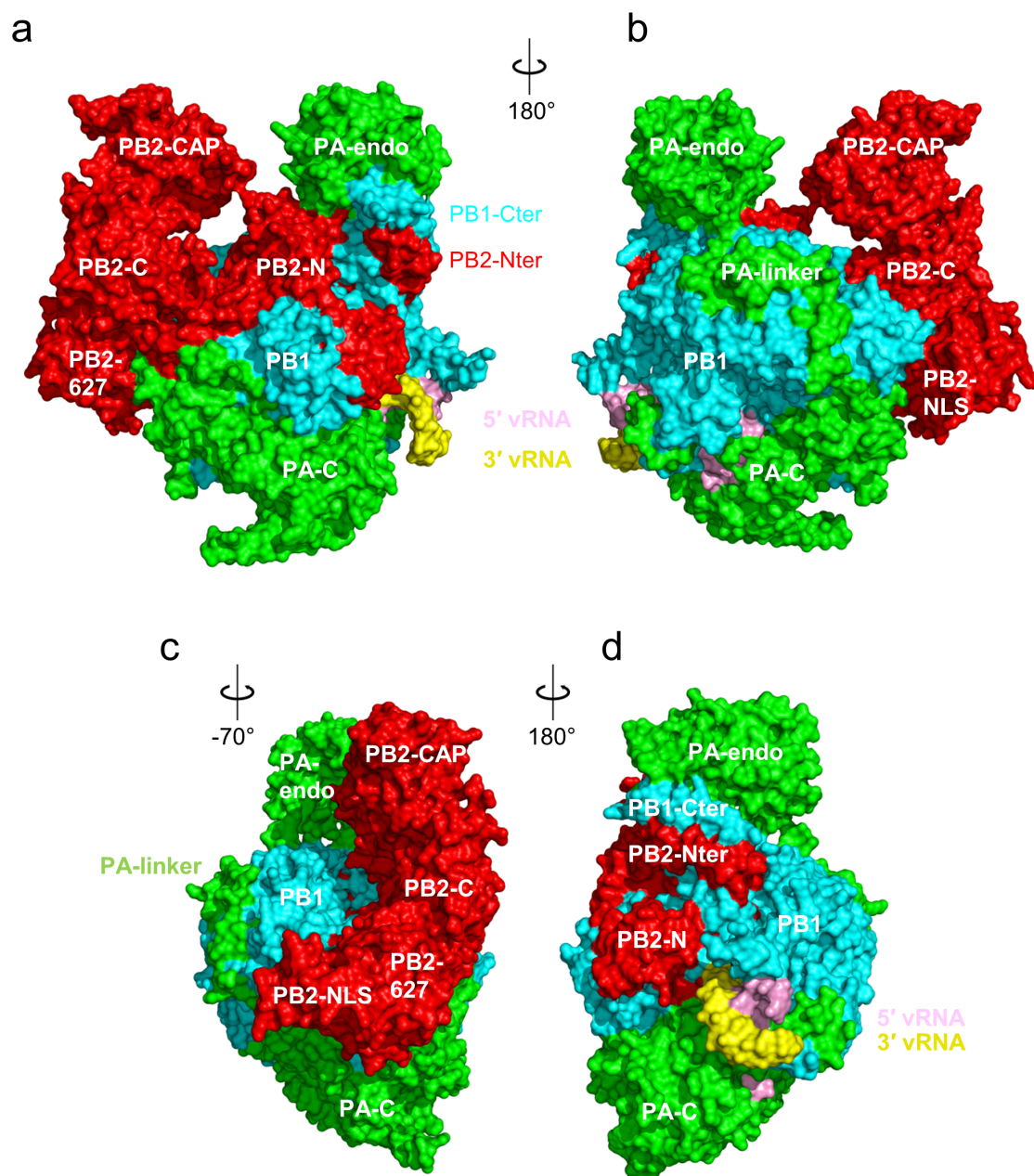
a, The heterotrimeric bat polymerase was recombinantly expressed in insect cells as a self-cleaving polyprotein. N-terminally it encodes the tobacco etch virus (TEV) protease that cleaves C-terminal to the amino-acid sequence ENLYFQ (in *italics*), and releases N-terminally His-tagged PA, PB1, C-terminally strep-tagged PB2 and cyan fluorescent protein (CFP) for facilitated monitoring of expression. Arrows indicate the N-to-C-terminal direction and the termini of each mature protein. The histidine and streptavidin tags are underlined. **b**, After ammonium sulphate precipitation, immobilized metal ion affinity chromatography, engineered streptavidin (strep-tactin) affinity and heparin chromatography, the final purification step consisted of

size-exclusion chromatography. The elution profile (monitored by the absorbance at 280 nm) with a single and nearly symmetric peak suggests a homogeneous and monomeric polymerase complex. mAU, milli-absorption unit. **c**, Fractions of the final size-exclusion chromatography were subjected to 10% SDS-PAGE followed by Coomassie blue staining. Lane 1 contains the molecular mass markers and lanes 2–7 the eluate with PA (85.4 kilodaltons (kDa)), PB1 (87.8 kDa) and PB2 (91.0 kDa). **d**, Recombinant bat FluA polymerase was visualized by electron microscopy following negative staining with sodium silico-tungstate of a 0.02 mg ml^{-1} protein sample. The image demonstrates that the sample is homogeneous and monodisperse with a V- or doughnut-like shape and central cavity.



Extended Data Figure 2 | Endonuclease, RNA transcription and RNA replication activities of recombinant FluA polymerase. **a**, Mini-panhandle vRNA: 5'-pppAGUAGUAACAAGAGGGUAUUGUAUACCUCUGCUUC UGCU-3'. **b**, Separate 5' and 3' ends: 5': 5'-pppAGUAGUAACAAGAGGGUAUUGUAUACCUCUGCUUC UGCU-3'; 3': 5'-UAUACCUCUGCUUCUGCU-3'. **c**, Endonuclease, cap-dependent transcription and ApG-primed replication assays. Cleavage of the cap donor is visible in lanes 2–6. Capped transcripts are visible in lanes 10 (from

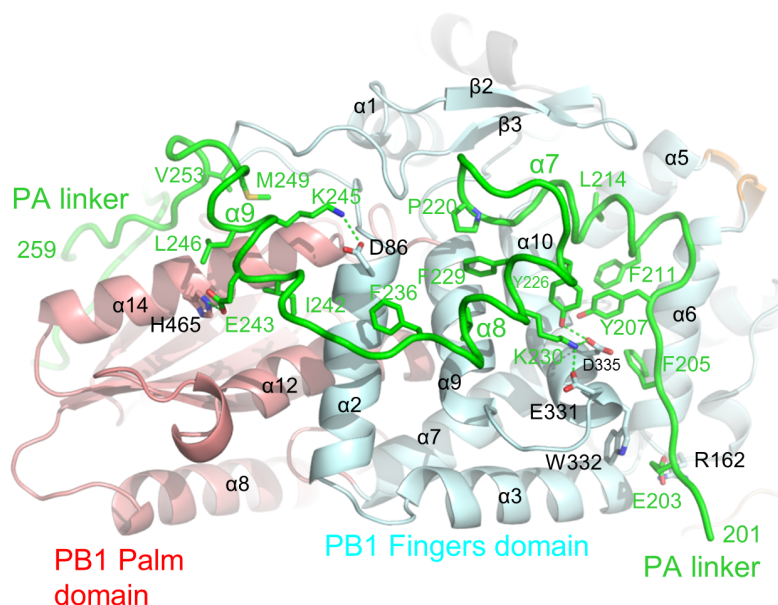
vRNA panhandle template) and 13 (from separated 5' and 3' vRNA ends) as well as cRNA produced in lanes 17 and 20. Markers, with size shown on the left, are RNA ladders labelled with ^{32}P -pCp nucleotide. **d**, **e**, Time course of unprimed (**d**) and ApG-primed (**e**) vRNA replication by bat influenza A polymerase. The products of replication (cRNA) are indicated with an arrow. Ladders (lanes L) are ^{32}P -pCp nucleotide-labelled RNA oligomers. ApG-primed replication is more efficient than unprimed replication.



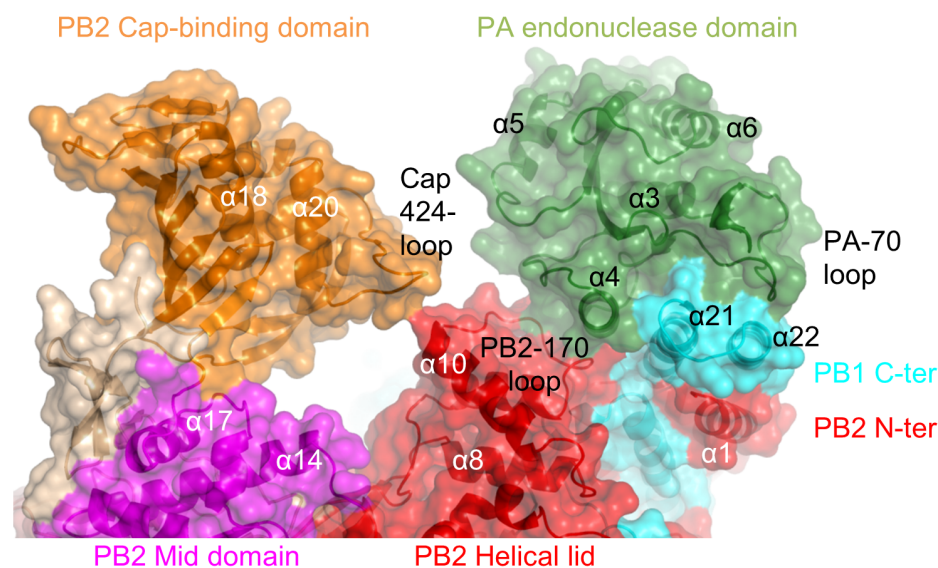
Extended Data Figure 3 | Surface views of the FluA heterotrimer with bound vRNA promoter. a–d, Four surface views at roughly 0° (a), 180° (b), 110° (c) and 290° (d) rotations with PA, PB1 and PB2 uniformly green,

cyan and red, respectively. Major subdomains are labelled. The vRNA 5' and 3' ends are pink and yellow, respectively.

a

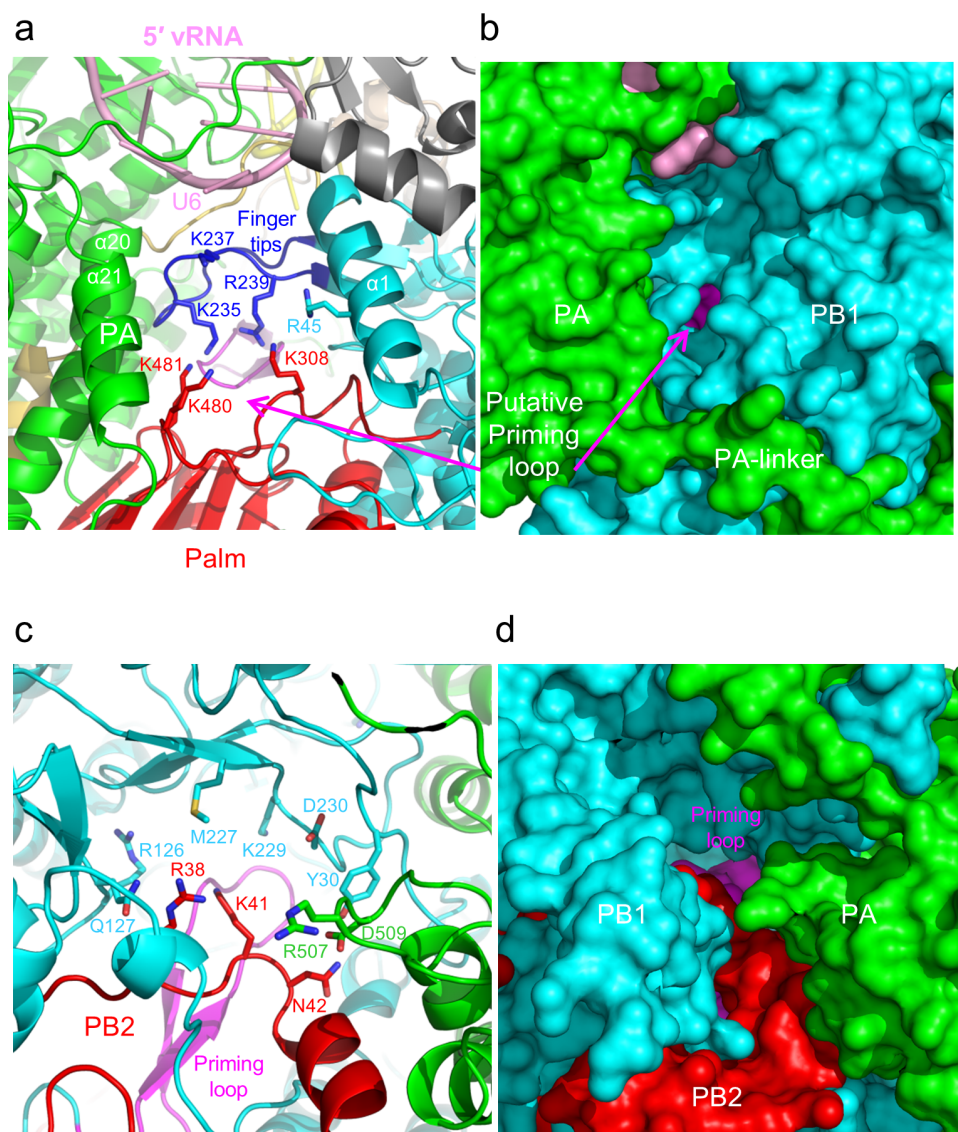


b



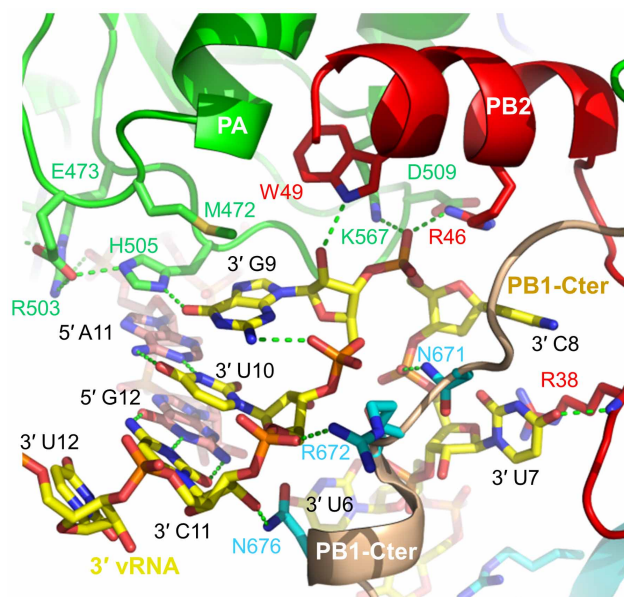
Extended Data Figure 4 | PA and PB2 structure and new inter-subunit interactions. **a**, Interactions of the PA-linker (green tube) with the outer surface of the fingers (pale cyan) and palm (pale salmon) domains of PB1. Contacts are mediated by both highly conserved hydrophobic residues (for example, PA residues Phe 205, Phe 211, Leu 214, Pro 220, Tyr 226, Phe 229, Tyr 232, Val 233, Ile 242, Leu 246, Met 249 and Val 253) and polar interactions (for example, PA Glu 203, Lys 230, Glu 243 and Lys 245 to PB1 Arg 162, Glu 331, His 465 and Asp 86, respectively). **b**, Transparent surface diagram showing the anchoring of the PA endonuclease domain (forest green) onto the

PB1-Cter–PB2-Nter interface region (cyan/red) and its position relative to the PB2 cap-binding domain (orange). The nuclease helix $\alpha 4$ packs parallel to the penultimate PB1 helix $\alpha 21$ involving both hydrophobic (for example, PA Ile 86, Ile 90 and Ile 94 with PB1 Ser 720, Ile 724 and Ile 728, respectively) and polar interactions (for example, PA Glu 77 with PB1 Arg 727). Other contacts include the PB2 170-loop interacting with the same PA helix $\alpha 4$ in the vicinity of Trp 88. Also the endonuclease insertion (PA 70-loop, residues 67–74) packs on the first part of the last PB1 helix $\alpha 22$. The total buried surface area between the endonuclease and PB1/PB2 is 2,265 Å².



Extended Data Figure 5 | NTP and template tunnels in PB1. **a**, View straight along the putative NTP entrance tunnel towards the putative priming loop (magenta) in the internal cavity. The NTP channel is lined with basic residues from the fingertips (Lys 235, Lys 237 and Arg 239, blue), fingers (Arg 45, cyan) and palm (Lys 308, Lys 480 and Lys 481, red) that are absolutely conserved in all influenza strains. The fingertips are in close proximity to PA helices $\alpha 20$ and $\alpha 21$ and to the loop of the 5' hook. **b**, Surface view as in **a** showing that the

putative priming loop in the interior cavity is visible through the NTP tunnel. **c**, View straight along the template entrance tunnel towards the priming loop (magenta) in the internal cavity. The tunnel is lined by residues conserved in all influenza strains and from all three subunits, Arg 507 and Asp 509 from PA (green), Tyr 30, Arg 126, Met 227, Lys 229 and Asp 230 from PB1 (cyan), and Arg 38, Lys 41 and Asn 42 from PB2 (red). **d**, Surface view as in **c** showing that the internal priming loop is visible through the template tunnel.



Extended Data Figure 6 | Recognition of the vRNA 3' end. Protein interactions of the distal 3' end showing the role of PB2-Nter (red). PB2 residues Arg 46 and Trp 49 and PA residue Lys 567 stabilize the sharp turn between 3' nucleotides C8 and G9. PB2 Arg 38 and PB1-Cter residues Asn 671, Arg 672 and Asn 676 also bind the 3' end. In the accompanying paper¹⁸, Fig. 2a shows the interactions with the complete 3' end as observed in the FluB vRNA complex.

Extended Data Table 1 | Data collection and refinement statistics for bat FluA polymerase

	Bat FluA
Data collection	
Space group	<i>C</i> 2
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	268.2, 149.3, 88.6,
α , β , γ (°)	90.0, 98.0, 90.0
Resolution (Å)	50.0-2.65(2.72-2.65)*
<i>R</i> _{meas}	8.2 (143.9)
<i>I</i> / σ <i>I</i>	10.87 (1.14)
Completeness (%)	99.2 (96.0)
Redundancy	3.5 (3.5)
Refinement	
Resolution (Å)	50.0-2.65 (2.71-2.65)
No. reflections	99462
<i>R</i> _{work} / <i>R</i> _{free}	21.5/26.7 (39.0/42.0)
No. atoms	
Protein	17517
RNA	600
Phosphates, zinc	65, 1
Water	46
B-factors (Å ²)	107.6
Protein	108.4
RNA	94.0
Water	74.9
R.m.s deviations	
Bond lengths (Å)	0.004
Bond angles (°)	0.732

*Highest resolution shell is shown in parenthesis.

Extended Data Table 2 | Direct polar polymerase–vRNA contacts for the bat FluA structure

PA (chain A) - 5' vRNA (chain V)									
Residue	Atom		RNA base	Atom				Distance (Å)	
Arg	279A	NH1	...	A	4V	OP1	...	3.42	*
Lys	281A	NZ	...	U	3V	O3'	...	3.37	*
			...	A	4V	OP1	...	3.17	***
His	326A	NE2	...	G	2V	OP1	...	2.96	***
Tyr	334A	OH	...	A	1V	O2'	...	3.50	*
Gly	366A	N	...	A	1V	N1	...	3.49	*
Gly	366A	O	...	A	10V	O2'	...	2.61	***
Gly	368A	N	...	A	11V	OP1	...	3.28	***
			...	A	11V	OP2	...	3.03	***
Gln	369A	N	...	A	11V	OP1	...	3.18	***
Ala	370A	N	...	A	11V	OP1	...	2.97	***
Tyr	388A	OH	...	C	9V	OP2	...	3.37	*
Lys	389A	O	...	A	7V	N7	...	3.06	***
Ser	390A	O	...	U	6V	O2	...	3.25	***
Arg	503A	NH1	...	A	11V	OP2	...	2.63	***
Arg	503A	NH2	...	A	11V	N7	...	3.34	*
His	505A	ND1	...	A	11V	N6	...	3.22	***
His	505A	NE2	...	A	11V	N7	...	3.19	***
Val	512A	N	...	C	9V	O2'	...	2.92	***
Asn	514A	OD1	...	G	2V	N3	...	3.43	*
			...	G	2V	N2	...	3.26	***
Asn	514A	ND2	...	A	10V	O2'	...	3.16	***
Lys	534A	NZ	...	U	3V	OP1	...	3.21	***
			...	U	3V	OP2	...	3.43	*
Arg	561A	NH2	...	G	2V	O3'	...	3.36	*
			...	U	3V	OP1	...	2.38	***
Thr	562A	O	...	G	2V	O2'	...	3.16	***
Lys	569A	NZ	...	A	4V	OP1	...	2.92	***
Asn	691A	ND2	...	G	5V	O6	...	3.47	*
PB1 (chain B) - 5' vRNA (chain V)									
His	32B	NE2	...	A	4V	O3'	...	3.13	***
			...	G	5V	OP1	...	2.88	***
Thr	34B	OG1	...	A	8V	OP1	...	2.81	***
Tyr	38B	OH	...	U	6V	O5'	...	3.08	***
Lys	188B	NZ	...	G	16V	O3'	...	3.21	***
			...	G	15V	OP1	...	3.35	*
Thr	201B	OG1	...	G	16V	OP2	...	3.31	*
			...	G	15V	O3'	...	2.57	***
Arg	203B	NE	...	G	14V	O2'	...	3.39	*
Arg	203B	NH1	...	G	14V	O4'	...	3.50	*
Met	356B	N	...	A	8V	OP2	...	3.31	*
Glu	358B	OE2	...	A	10V	N6	...	3.32	*
Arg	365B	NE	...	C	9V	OP2	...	3.27	***
Arg	365B	NH2	...	A	10V	OP2	...	2.98	***
			...	A	10V	N7	...	3.15	***
PA (chain A) - 3' vRNA (chain T, numbered from 3' end)									
His	505A	ND1	...	G	9T	O6	...	2.80	***
Arg	507A	NE	...	U	7T	OP2	...	2.97	***
Arg	507A	NH1	...	U	10T	O2	...	3.23	***
Arg	507A	NH2	...	U	10T	O2'	...	2.43	***
			...	U	10T	O2	...	3.04	***
Lys	567A	NZ	...	C	8T	OP2	...	2.92	***
PB1 (chain B) - 3' vRNA (chain T, numbered from 3' end)									
Asn	671B	ND2	...	U	7T	OP1	...	2.86	***
Arg	672B	N	...	G	9T	OP1	...	2.85	***
Arg	672B	NH1	...	U	10T	OP1	...	2.88	***
Ser	673B	N	...	U	10T	O3'	...	3.43	*
Ser	673B	OG	...	U	10T	O2'	...	2.99	***
			...	U	7T	OP1	...	2.46	***
Ile	674B	N	...	U	7T	O2'	...	3.39	*
Asn	676B	OD1	...	C	11T	O2'	...	2.58	***
Asn	676B	ND2	...	U	10T	O4'	...	3.14	***
PB2(chain C) - 3' vRNA (chain T, numbered from 3' end)									
Ser	36C	O	...	U	7T	N3	...	3.11	***
Arg	38C	N	...	U	7T	O4	...	2.73	***
Arg	38C	NH1	...	U	7T	O4'	...	3.40	*
			...	C	8T	O2	...	3.33	*
Arg	38C	NH2	...	U	7T	O4'	...	2.69	***
Glu	40C	OE1	...	C	8T	N4	...	3.39	*
Arg	46C	NH1	...	C	8T	OP2	...	2.54	***
Arg	46C	NH2	...	C	8T	O4'	...	3.27	***
Trp	49C	NE1	...	G	9T	O2'	...	2.70	***

Structural insight into cap-snatching and RNA synthesis by influenza polymerase

Stefan Reich^{1,2*}, Delphine Guigligay^{1,2*}, Alexander Pflug^{1,2*}, Hélène Malet^{1,2}, Imre Berger^{1,2}, Thibaut Crépin², Darren Hart^{1,2}, Thomas Lunardi^{1,2}, Max Nanao^{1,2}, Rob W. H. Ruigrok² & Stephen Cusack^{1,2}

Influenza virus polymerase uses a capped primer, derived by ‘cap-snatching’ from host pre-messenger RNA, to transcribe its RNA genome into mRNA and a stuttering mechanism to generate the poly(A) tail. By contrast, genome replication is unprimed and generates exact full-length copies of the template. Here we use crystal structures of bat influenza A and human influenza B polymerases (FluA and FluB), bound to the viral RNA promoter, to give mechanistic insight into these distinct processes. In the FluA structure, a loop analogous to the priming loop of flavivirus polymerases suggests that influenza could initiate unprimed template replication by a similar mechanism. Comparing the FluA and FluB structures suggests that cap-snatching involves *in situ* rotation of the PB2 cap-binding domain to direct the capped primer first towards the endonuclease and then into the polymerase active site. The polymerase probably undergoes considerable conformational changes to convert the observed pre-initiation state into the active initiation and elongation states.

The influenza virus genome comprises eight segments of single-stranded RNA (viral RNA, or vRNA), each packaged in separate ribonucleoprotein particles (RNPs). Both conserved 3' and 5' ends of the vRNA (the promoter) are bound to the RNA-dependent RNA polymerase, and the rest of the pseudo-circularised vRNA is coated with nucleoprotein. The polymerase is a heterotrimer composed of subunits PA, PB1 and PB2 and, in the context of the RNP, it performs the distinct processes of transcription and replication using the same template vRNA (reviewed in refs 1 and 2). Transcription of viral mRNA occurs through a unique process called cap-snatching, in which short capped oligomers, derived from host pre-mRNA, are bound by the PB2 subunit^{3–5}, cleaved by an endonuclease in the PA subunit^{6,7} and then used to prime mRNA synthesis by the PB1 subunit. Stuttering of the polymerase at an oligo-U stretch near the vRNA 5' end leads to auto-polyadenylation⁸. Thus, translation-competent viral mRNAs are generated without the need for a viral-encoded capping enzyme nor the host poly-adenylation machinery, which is shut down by viral-encoded NS1 protein⁹. By contrast, replication involves unprimed synthesis of an exact, full-length copy of the vRNA into complementary RNA (cRNA) and subsequently the inverse process back to progeny vRNA. Nascent replicates are immediately packaged with nucleoprotein into new viral RNPs (vRNPs) or complementary RNPs (cRNPs). In contrast, viral mRNA is not so packaged but is treated as host pre-mRNA¹⁰ and further spliced (in the case of NS and M segments) and/or exported to the cytoplasm by host cell machineries. Interestingly, cRNPs do not perform transcription in infected cells and may require a second polymerase to replicate^{11,12}. Despite many years of study, the mechanism by which RNPs are able to perform these different functions and what determines the type of RNA synthesis that occurs are still obscure. Here we infer, using complementary information from atomic resolution structures of influenza A and B polymerases in complex with the vRNA promoter together with known structures of other viral RNA polymerases, the mechanism by which the polymerase can perform either cap-dependent transcription or unprimed (*de novo*) RNA synthesis. The structures thus open the way to a detailed description of how the

influenza transcription/replication machine works in a context-dependent manner.

Structure of FluB polymerase compared with FluA

Full-length heterotrimeric influenza polymerase from B/Memphis/13/03 (FluB) was obtained by expression in insect cells as a self-cleaving polypeptide¹³ (Extended Data Fig. 1). Recombinant FluB polymerase was active in the absence of nucleoprotein in cap-dependent transcription and both ApG-primed and, less efficiently, unprimed replication assays using short model vRNAs (Extended Data Fig. 2). Two different crystal forms of FluB polymerase were obtained with consensus promoter sequences for influenza B¹⁴ (Extended Data Table 1). Both contain nucleotides 1–14 from the vRNA 5' end (5'-pAGUAGUACAAGAG-3'OH) and either nucleotides 5–18 (FluB1 form) or nucleotides 1–18 (FluB2 form) from the 3' end (3'OH-UCGUCUUCGUCUCCAUAU-5'OH). The FluB1 form yielded a fully interpretable experimental map (Extended Data Fig. 3a–c) at 3.4 Å resolution, allowing an almost complete model of FluB polymerase to be built (Fig. 1a). The 2.7-Å resolution FluB2 structure, solved by molecular replacement using the FluB1 structure, is extremely well ordered (Extended Data Fig. 3d, e). Owing to crystal contacts, it has the best defined endonuclease domain, which, however, is in the same position as in all other structures. By contrast, the C-terminal two-thirds of PB2 (PB2-C) completely lacks electron density in the FluB2 form (Fig. 1b), although intact PB2 is present in the crystal.

Sequence alignments with bat FluA, the structure of which is described elsewhere¹⁵, show that influenza B/Memphis/13/03 has 36.0 (48.6), 59.5 (71.0) and 37.0 (50.9) per cent amino acid identity (similarity) for PA, PB1 and PB2, respectively (with higher than average conservation in the functionally important regions) (Supplementary Fig. 1). The FluA and FluB polymerase structures and their mode of binding to the vRNA promoter are remarkably similar (Extended Data Fig. 4). However, a striking difference of 70° in the orientation of the PB2 cap-binding domain (Fig. 1c, d) suggests that this domain can rotate *in situ*. Concerning

¹European Molecular Biology Laboratory, Grenoble Outstation, 71 Avenue des Martyrs, CS 90181, 38042 Grenoble Cedex 9, France. ²University Grenoble Alpes-Centre National de la Recherche Scientifique-EMBL Unit of Virus Host-Cell Interactions, 71 Avenue des Martyrs, CS 90181, 38042 Grenoble Cedex 9, France.

*These authors contributed equally to this work.

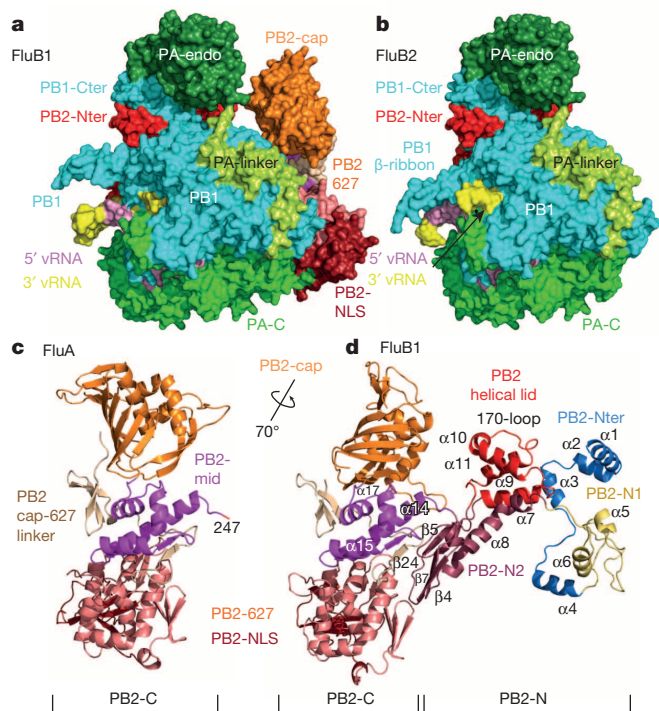


Figure 1 | Structure of influenza B polymerase. **a**, Surface view of FluB1 structure colour-coded according to domain structure (Extended Data Fig. 4) except that PA-C, PB1 and PB2-N are uniformly green, cyan and red, respectively. The vRNA 5' and 3' ends are pink and yellow, respectively. **b**, Surface view of the polymerase in the FluB2 crystal form that lacks the entire PB2-C domain but includes the full-length 3' end of the vRNA (black arrow). **c**, Bat FluA PB2-C colour-coded according to domain structure (Extended Data Fig. 4). **d**, The complete PB2 subunit as in the FluB1 crystal form in the same orientation as in **c**, highlighting the 70° difference in orientation of the cap-binding domain.

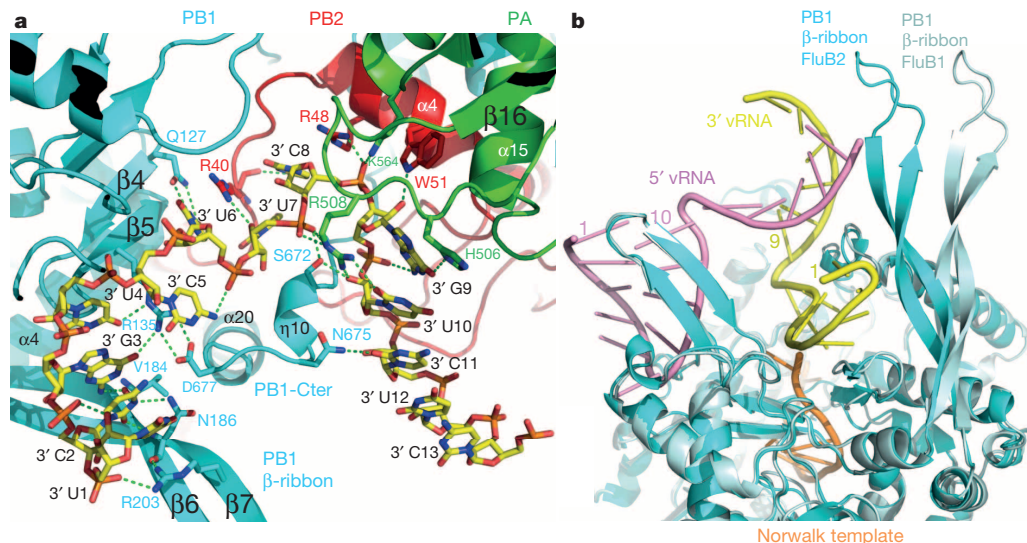


Figure 2 | Promoter 3'-end binding and PB1 β-ribbon flexibility. **a**, Diagram showing RNA-RNA and RNA-protein interactions of the complete 3' end (nucleotides 1–13, yellow sticks) of the promoter as in the FluB2 structure. For clarity, not all interactions (nor water-mediated interactions) are depicted. All three subunits, PA (green ribbons and residues), PB1 (cyan) and PB2 (red), are involved. Nucleotides 1–9 are single-stranded, and 10–13 form a duplex with the 5' end (not shown). The PB1 β-ribbon interacts with the proximal part of the 3' end and PB1-Cter interacts with

vRNA binding, all FluA and FluB structures exhibit identical conformations of 5' nucleotides 1–14 and 3' nucleotides 5–13 as described elsewhere¹⁵, and most protein–RNA contacts for these regions are conserved between FluA and FluB (Extended Data Table 2). The higher resolution FluB2 structure shows that the protein–RNA interface is highly hydrated with numerous water-mediated protein–RNA interactions. In both the FluB1 and FluB2 structures, the 3'–5' duplex region of the promoter comprises four base pairs (3' 10-UCUC-13 with 5' 11-AG AG-14). In the FluB1 structure, the 5-nucleotide 3' overhang (14-CA UAU-18) forms a triple-stranded structure at a two-fold crystal contact, including two base triples with the symmetry-related duplex (Extended Data Fig. 5), whereas in the FluB2 structure the RNA does not participate in crystal contacts.

Promoter 3' end binding

Only in the FluB2 crystal form is the complete 3' end of the promoter structurally ordered (including nucleotides 1–5, not visible in other structures). The single-stranded 3' extremity, 1-UCGUCUUCG-9, perhaps unexpectedly, does not enter the polymerase active site but is bound in an alternative location on the surface of the polymerase in an arc conformation, such that U1 is not far from the distal 3'–5' duplex region (Fig. 2a). Bases 1–4 stack on each other but other bases are bound in individual pockets. Most bases are orientated towards the protein and all except U1 make base-specific RNA–RNA or RNA–protein interactions (Fig. 2a and Extended Data Table 2). All three subunits are involved in binding 3' nucleotides 6–9, whereas nucleotides 1–5 only interact with PB1. Residues 670–679 of PB1 are involved in binding both extremities of the 3' end whereas the PB1 β-ribbon interacts with 3' nucleotides 1–3 (Fig. 2a). The sequence-specific nature of the 3'-end binding and conservation of interacting residues strongly suggests that this binding site is functionally important. This implies that there must be a mechanism for relocating the 3' end into the PB1 active site during initiation of RNA synthesis. The observed 3'-end conformation is inconsistent with a hook conformation¹⁶, but overall the promoter structure is consistent with that proposed in ref. 17, which suggests that the sequence constraints imposed on the 3' end by the necessity of almost exact complementary to the 5' hook would make it appear that the 3' end would also take a hook conformation.

both proximal and distal nucleotides. Specific RNA–RNA interactions include N2 to OP2 of G9, N4 of C5 to OP2 of U6, O2' of U1 to OP2 of C2. **b**, Superposition of PB1 from the FluB2 (cyan) and FluB1 (light cyan) structures, showing flexibility of the long β-ribbon. The 3' (nucleotides 1–13, yellow) and 5' (nucleotides 1–14, pink) ends of the promoter are as in the FluB2 structure. The 3' deviates from the path into the PB1 active site that is depicted by the template strand (orange) from the superposed Norwalk template–primer elongation complex (PDB code 3BSO).

In all structures, the unusually long PB1 β -ribbon (residues 177–214) has a role in interacting with the vRNA on the exterior of the polymerase. In the FluB1 structure, the β -ribbon is straight and projects away from the polymerase, its tip (residues 195–196) interacting with crystal symmetry-related RNA (Fig. 2b and Extended Data Fig. 5). In the bat FluA structure, the ribbon is bent towards the polymerase and its central part contacts the duplex region of the promoter, whereas its extremity is disordered (not shown). In the FluB2 structure, the β -ribbon is the most bent and residues 184–186 and influenza-conserved Arg 203 interact with the proximal 3' end (Fig. 2a, b). These observations show that the PB1 β -ribbon has an affinity for RNA and is flexible. It could therefore have a dynamic role in translocating the RNA into the polymerase from the RNP and/or could mediate interactions with proximal nucleoprotein molecules of the RNP. This hypothesis is supported by fitting of the polymerase–promoter structure to the available electron microscopy map of the mini-RNPs, which predicts the close proximity of the ribbon to nucleoprotein (Extended Data Fig. 6).

Mechanism of replication

Influenza virus polymerase catalyses primer-independent (*de novo*) replication to generate cRNA from vRNA and vice versa. It has been proposed that efficient replication requires nucleoprotein¹⁸ and/or polymerase oligomers^{11,12,19}. RNA polymerases that perform *de novo* synthesis generally possess a special 'priming' loop that is thought to stabilize the priming and incoming NTPs in the absence of a priming oligonucleotide. This phenomenon was first structurally characterized for bacteriophage $\Phi 6$ polymerase, in which a tyrosine at the extremity of the priming loop stacks on the priming nucleotide²⁰. Flavivirus polymerases such as those of hepatitis C virus (HCV) or dengue virus (DENV) also have an aromatic residue (a tyrosine in HCV and a tryptophan for DENV) as putative priming platforms²¹. For PB1, a β -hairpin loop (residues 641–657), structurally analogous to that of HCV, is observed in an ordered configuration in the FluA structure¹⁵ but is disordered in the FluB polymerase structures. The loop tip contains the 648-Ala-His-Gly-Pro motif,

conserved in all influenza polymerases. Modelling, on the basis of the $\Phi 6$ initiation complex structure, shows that the loop could potentially act as a priming platform to promote correct initiation, with His 649 plausibly interacting with the initial incoming nucleotides (Fig. 3a, b). More details of the active site configuration, which largely involves the canonical polymerase motifs, are given for the FluA structure¹⁵. A model of the elongation step of influenza polymerase can be obtained by superposing the primer–template complex of poliovirus polymerase²² on PB1, the high conservation of the polymerase active site ensuring an unambiguous superposition (Fig. 3c, d). The putative priming loop would need to be displaced once elongation starts because it would sterically clash with an emerging template–product duplex (Fig. 3c).

These results lead to the following two observations. As highlighted above, in our structures, the vRNA 3' extremity does not enter the PB1 active site. However, comparison for instance with the polio template–product complex shows that vRNA 3' nucleotide 8/9 is at the template tunnel entrance but corresponds to 3' nucleotide 5/6 in the polio polymerase complex (Fig. 3d). Thus, the 3' end, on reorientation into the PB1 active site, would have to draw back three nucleotides to initiate at the first position, perhaps concomitantly with breaking of the 3'–5' duplex region. The mechanism to do this is unclear at present. Interestingly, it has been proposed that vRNA and cRNA initiate replication differently, either synthesizing pppApG at positions 1 and 2 (1-UC) of the 3' end directly for vRNA, or internally at 4-UC followed by realignment at 1-UC for cRNA²³. This suggests that the 3' end is differently positioned in the active site depending on whether it is vRNA or cRNA. This could be because the c3' end sequence differs at three positions and is one nucleotide longer than the v3' end before the 3'–5' duplex region. According to the modelling, it would thus be positioned correctly for internal initiation and the putative priming loop could have a role in this.

The second observation concerns the fact that modelling with the polio template–product elongation complex shows that an extended duplex cannot be accommodated in the cavity of the current structures

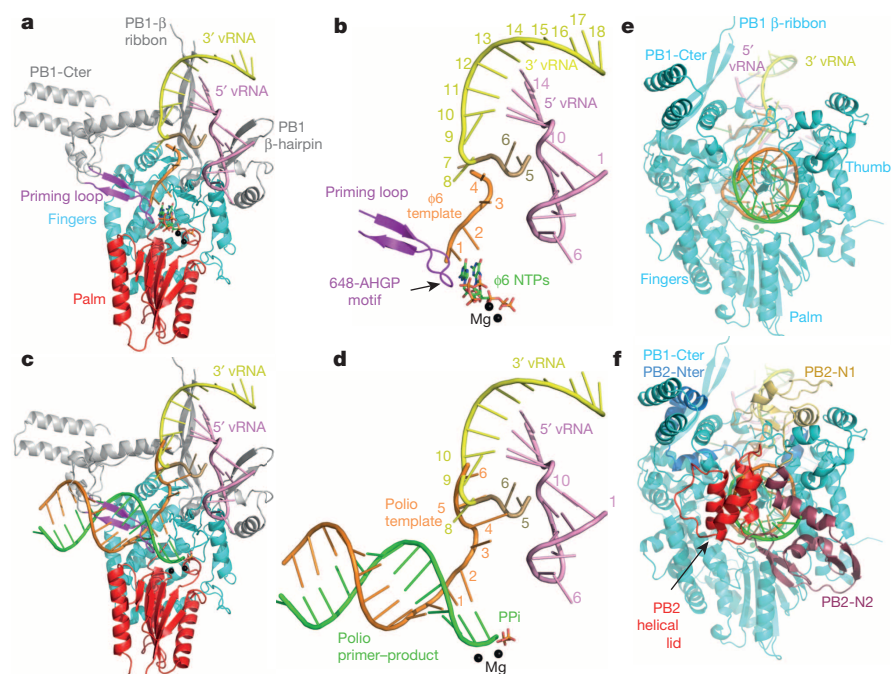


Figure 3 | Model for replication initiation and elongation by influenza polymerase. **a**, FluA PB1 with bound 3' (nucleotides 5–18, yellow) and 5' (nucleotides 1–14, pink) vRNA superposed with the $\Phi 6$ initiation complex structure (PDB code 1HI0) with template (orange) and two initial incoming NTPs with magnesium (green sticks and black spheres). The PB1 putative priming loop is magenta with the palm (red) and fingers (cyan). The thumb is omitted for clarity. **b**, As in **a** but showing only the PB1 putative priming loop. The influenza conserved 648-Ala-His-Gly-Pro motif at the loop tip could stabilize the initiation complex (the electron density for the His-Gly residues is poor). 3'-end nucleotides 5 and 6 that deviate from the canonical template pathway (orange) are sand coloured. **c**, As in **a** but with the primer (green) and template (orange) RNA from the poliovirus polymerase elongation complex (PDB code 3OL7) after superposition of the polymerase domains. The PB1 putative priming loop clashes with the duplex RNA and therefore must be displaced. **d**, As in **c** but excluding the protein. The influenza vRNA 3' nucleotide 8/9 is at the template entrance but corresponds to the 3' nucleotide 5/6 template in the polio virus polymerase complex. **e**, As in **c** but end-on view and with PB1 uniformly coloured cyan. The template–product duplex can be accommodated in the PB1 cavity although the thumb domain is expected to open. **f**, As in **e** but including the PB2-N domain with subdomains coloured as in Extended Data Fig. 4. The product/ primer strand (green) can potentially exit/enter (see also Fig. 5) but the template strand is blocked by the PB2 helical lid domain (red).

because of a severe steric clash of the outgoing template strand. This is not primarily due to PB1 (apart from the putative priming loop) (Fig. 3e), but to elements of PB2-N that lie directly on top of the duplex (Fig. 3f). As discussed below, the current structures do provide an open channel into the PB1 active site for a capped primer to initiate transcription but the outgoing template is blocked by helices $\alpha 8$ – $\alpha 10$ of the PB2 lid domain. In the case of HCV polymerase, the structure of a product–template duplex complex revealed a 20° rotation of the thumb domain that opened up the product–template binding cavity²⁴. In analogy to this, we expect elongation to be accompanied by equivalent conformational changes in which thumb opening could be coupled to displacement of the priming loop and rotation out of the way of the PB2 lid domain. The length of the product–template duplex that is accommodated by influenza polymerase, what causes strand separation (although the PB2 N2 and lid domains are plausibly involved) and which exit path the two strands take are open questions. As discussed below, in the case of cap-dependent transcription, a likely exit pathway for the nascent mRNA is away from the nuclease domain and towards the PB2 627-domain (containing the host-specific amino acid residue 627).

Mechanism of cap-dependent transcription

Cap-snatching is uniquely performed by segmented negative-strand RNA viruses including orthomyxoviruses, bunyaviruses and arenaviruses^{25,26}. The PB2 cap-binding and PA endonuclease domains involved in this process were previously characterized structurally and functionally^{5–7}. The complete polymerase structures now allow a plausible mechanism for cap-snatching and cap-dependent priming to be proposed. All structures show the PA-Nter endonuclease in the same position and orientation, anchored to the PB1-Cter–PB2-Nter interface (Fig. 1a). By contrast, comparison of the FluA and FluB1 structures after superposition suggests that the PB2 cap-binding domain is able to rotate as a rigid body *in situ*. Whereas PB2 residues before Ile 319 (Ile 321 in FluB) and after Arg 495 (Lys 496 in FluB) align very well, the entire cap-binding domain in between differs in orientation by 70° between the two structures, suggesting that it is flexibly hinged at these anchor points (Fig. 1c, d and Supplementary Video 1). In the FluA structure, the cap-binding site faces the endonuclease active site directly across a solvent channel at a distance of about 50 Å (Fig. 4a, b). This configuration is consistent with a cap-bound host pre-mRNA being cleaved 10–15 nucleotides downstream by the nuclease²⁷, bearing in mind that the observed cap-binding domain orientation, probably constrained by crystal contacts, is not necessarily optimal for cap-snatching. The observed variability of the primer length²⁷ would be explained by flexibility in both the cap-binding domain orientation and RNA conformation and possibly the sequence preference of the nuclease cleavage site²⁸. Cleaved primers would then be further selected by their efficiency in priming mRNA synthesis, which probably correlates with the complementary to the extremity of the 3' template^{27,29–32}. In the FluB1 structure, the rotated position of the cap-binding domain both shields the bound-capped primer from the endonuclease (Fig. 4c, d) and directs it down into the polymerase RNA catalytic cavity (Fig. 4e). This model is supported by the observation in FluB1 crystals (but not other crystal forms) of residual difference electron density, strongly suggestive of RNA, that descends precisely from the Trp 369–Phe 406 sandwich in the FluB cap-binding site into the throat of the polymerase, which leads to the PB1 active site (Fig. 4e and Extended Data Fig. 7). The nature and origin of this RNA is unclear, making it difficult to fit a precise model. But whatever the RNA origin, its fortuitous occurrence in the FluB1 structure gives a very plausible model of how a capped primer might be configured during transcription initiation. The 424-loop of the cap-binding domain seems to have key roles in channelling the capped primer into the polymerase throat (the integrity of this loop was previously shown to be important for transcription⁵), as well as the projecting amino-terminal end of PB2 lid domain helix $\alpha 9$, and in particular the double prolines 157–Pro–Pro that force the RNA into a $\sim 90^\circ$ bend (Fig. 4e and Extended Data Fig. 7). The observed RNA density corresponds to about six nucleotides plus

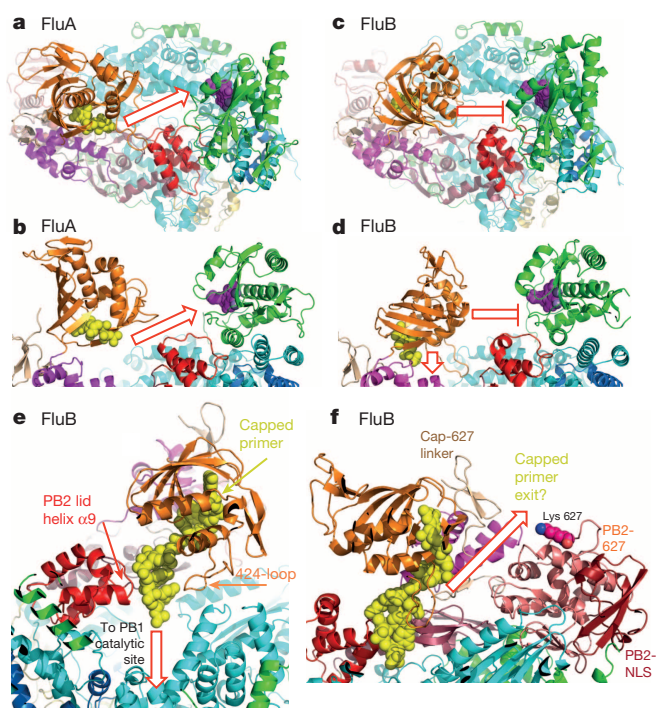


Figure 4 | Cap-snatching and cap-dependent priming of transcription.

a, Cap-snatching configuration. Top view of the relative orientations in the FluA structure of the cap-binding domain (orange) with bound cap analogue (m^7 GTP, yellow spheres, obtained by superposition with PDB code 4CB4) and endonuclease (green) with active site indicated by a bound inhibitor (purple spheres). PA is uniformly green, PB1 cyan and PB2 subdomains coloured as in Extended Data Fig. 4. Cap-bound host pre-mRNA can reach the endonuclease active site unimpeded across a solvent channel (red arrow) at a straight-line distance of around 50 Å (although the cap-binding domain orientation observed is not necessarily optimal for primer cleavage). **b**, As in **a** but side view. **c**, **d**, As in **a** and **b** but for the FluB1 structure. The rotated orientation of the cap-binding domain shields cap-bound RNA from the endonuclease. **e**, The FluB1 cap-binding domain configuration is compatible with cap-dependent priming. Yellow spheres represent model of bound capped primer derived from RNA-like residual electron density (Extended Data Fig. 7). The primer is channelled towards the PB1 active site by the 424-loop of the cap-binding domain and the N-terminal end of PB2 lid domain helix $\alpha 9$. **f**, The putative exit channel for the capped transcript is between the PB2 cap-627 linker and 627-domains towards host-specific residue Lys 627, and away from the nuclease.

the cap, and extends over a straight-line distance of around 26 Å to the bend. The remaining distance to the polymerase active site is around 28 Å, which is compatible with a primer of around 12–14 nucleotides (Extended Data Fig. 7). An overall model of how cap-dependent priming is likely to occur in influenza polymerase is given in Fig. 5.

One can hypothesize about subsequent steps in cap-dependent transcription (Extended Data Fig. 8). Once the capped primer 3' end engages the vRNA template in the PB1 active site, primer elongation occurs by template-directed nucleotide addition. Further rotation *in situ* of the cap-binding domain could initially accommodate the growing mRNA while still maintaining cap-binding (Extended Data Fig. 8c). However, at some stage the buckling out of the lengthening mRNA would force cap release, which has previously been estimated to be after 11–15 nucleotides³³ (Extended Data Fig. 8d). The transcript would naturally emerge into the basic channel between the cap-binding domain and the cap-627 linker/627-domain in the vicinity of host-specific residue 627 (Fig. 4f), possibly explaining why capped RNA was crosslinked to the 627-domain^{34,35}. This exit pathway avoids the endonuclease, consistent with reports that the polymerase protects its own mRNAs from degradation by transiently binding to the conserved AGCAAAGCAGG

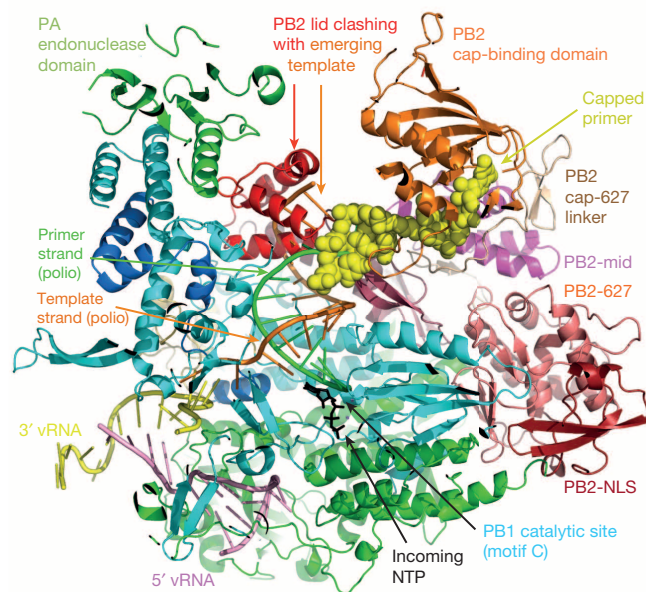


Figure 5 | Model for cap-dependent transcription. The FluB1 structure is superposed with the template–primer (orange/green) duplex and incoming NTP (black sticks) from the poliovirus complex (PDB code 3OL7). PA is uniformly green, PB1 cyan and PB2 subdomains coloured as in Extended Data Fig. 4. The capped RNA primer (yellow spheres) is as in Fig. 4e and connects with the primer strand in the polio complex. The polio template strand connects with the vRNA 3' end (yellow tube) at the template tunnel entrance. During elongation the emerging template strand would clash with the PB2 helical lid, which therefore has to move. When the vRNA template has mostly passed through the polymerase there will be a minimal loop remaining with the tightly bound 5' hook (pink tube), which will generate the poly(A) tail on the transcript by stuttering on the oligo-U sequence at 17–22 nucleotides from the 5' end. During elongation, the polymerase will sequester 20–25 nucleotides of the template. In the context of transcription by an RNP, at least this amount of RNA would have to dissociate from nucleoprotein and re-associate after exiting the polymerase.

sequence, which occurs just downstream of the host mRNA-derived primer sequence and is transcribed from the conserved 3' end of the template³⁶. The 627-domain may have a role in this as it has a binding preference for 5' vRNA-like sequences³⁷. When eventually released, the 5' cap structure itself is bound by the nuclear cap-binding complex and the mRNA subsequently processed by host machinery¹⁰ (Extended Data Fig. 8d). More generally, the same exit pathway could be used for cRNA or vRNA replicates, and the 627-NLS domain (a double domain containing host-specific PB2 residue 627 and the PB2 nuclear localization signal (NLS)) could have a role in their packaging with incoming nucleoprotein into nascent cRNPs or vRNPs³⁸.

Concerning auto-polyadenylation of viral mRNA by the polymerase, the tight binding of the hook at the 5' end of the template is thought to cause stuttering at the oligo-U stretch typically 17–22 nucleotides from the 5' end, resulting in the addition of several adenosine residues^{8,17}. Because a minimum of ten 5' nucleotides are required to form the 5' hook, and, on the basis of the structural alignment with the polio polymerase primer–template complex, a minimum of seven extra nucleotides is required to reach the site of nucleotide addition (Fig. 3e), the crystal structure is fully compatible with the proposed polyadenylation model (Extended Data Fig. 8d).

Conclusions

The FluA and FluB polymerase structures presented here seem to be in an inactive pre-initiation state requiring relocation of the 3' end into the PB1 active site before RNA synthesis can begin. However, we think the observed 3'-end binding site on the polymerase surface could have functional importance in, for instance, providing an additional docking site for the 3' end (on the same polymerase or a different polymerase) after it has been copied and exited from the active site. This would be an efficient way to allow several rounds of primary transcription from the same vRNP in the early stages of infection. Alternatively, the 3' end bound to the surface of one polymerase could translocate into the active site of a second, empty polymerase as has been proposed in some models of replication that imply polymerase oligomerization¹¹. It is also clear, as discussed above, that additional conformational changes must occur to allow progression from the initiation to the elongation stages of RNA synthesis. Although it is expected, in analogy to HCV, that the PB1 RNA-binding cavity widens during this step and the lid made by the PB2 N2 domain should open (see above), several other lines of evidence suggest that PB2 as a whole is the most mobile part of the polymerase. First, docking of the polymerase crystal structure into the mini-RNP electron microscopy map³⁹ shows that the PA–PB1 heterodimer fits well but the extra density assigned to PB2 is detached from the rest and cannot be fitted without a gross conformational change of PB2 (Extended Data Fig. 6). Second, detachment of a large fragment of PB2 is compatible with the polymerase structure in the FluB2 crystal form, in which two-thirds of PB2 (PB2-C) is not visible at all although there is space in the crystal for it. Similarly, in the electron microscopy reconstructions of native RNPs¹⁹, part of the polymerase (the 'arm'), was observed to be detached and flexible. Although this was assigned to PA-C, we think it is most likely to be PB2-C, for the reasons just given and because our structures suggest that the integrity of the PA–PB1 heterodimer is very unlikely to be disrupted at least while both are intimately binding the vRNA 5' end.

Although there are still very many open questions, our three complementary structures already give considerable new insight into the mechanism of replication and transcription by influenza polymerase. They provide a solid structural framework for future studies aimed at refining understanding of this complex and dynamic molecular machine, not only in isolation but also in the more complicated physiological context of the RNP and host factors.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 18 August; accepted 29 October 2014.

Published online 19 November 2014.

1. Resa-Infante, P., Jorba, N., Coloma, R. & Ortin, J. The influenza virus RNA synthesis machine: advances in its structure and function. *RNA Biol.* **8**, 207–215 (2011).
2. Fodor, E. The RNA polymerase of influenza A virus: mechanisms of viral transcription and replication. *Acta Virol.* **57**, 113–122 (2013).
3. Ulmanen, I., Broni, B. A. & Krug, R. M. Role of two of the influenza virus core P proteins in recognizing cap 1 structures (m⁷GpppNm) on RNAs and in initiating viral RNA transcription. *Proc. Natl Acad. Sci. USA* **78**, 7355–7359 (1981).
4. Blass, D., Patzelt, E. & Kuechler, E. Identification of the cap binding protein of influenza virus. *Nucleic Acids Res.* **10**, 4803–4812 (1982).
5. Guilligay, D. et al. The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nature Struct. Mol. Biol.* **15**, 500–506 (2008).
6. Dias, A. et al. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **458**, 914–918 (2009).
7. Yuan, P. et al. Crystal structure of an avian influenza polymerase PA_N reveals an endonuclease active site. *Nature* **458**, 909–913 (2009).
8. Pritlove, D. C., Poon, L. L., Fodor, E., Sharps, J. & Brownlee, G. G. Polyadenylation of influenza virus mRNA transcribed *in vitro* from model virion RNA templates: requirement for 5' conserved sequences. *J. Virol.* **72**, 1280–1286 (1998).
9. Nemeroff, M. E., Barabino, S. M., Li, Y., Keller, W. & Krug, R. M. Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3' end formation of cellular pre-mRNAs. *Mol. Cell* **1**, 991–1000 (1998).
10. Bier, K., York, A. & Fodor, E. Cellular cap-binding proteins associate with influenza virus mRNAs. *J. Gen. Virol.* **92**, 1627–1634 (2011).

11. Jorba, N., Coloma, R. & Ortin, J. Genetic trans-complementation establishes a new model for influenza virus RNA transcription and replication. *PLoS Pathog.* **5**, e1000462 (2009).
12. York, A., Hengrung, N., Vreede, F. T., Huiskonen, J. T. & Fodor, E. Isolation and characterization of the positive-sense replicative intermediate of a negative-strand RNA virus. *Proc. Natl Acad. Sci. USA* **110**, E4238–E4245 (2013).
13. Nie, Y., Bellon-Echeverria, I., Trowitzsch, S., Bieniossek, C. & Berger, I. Multiprotein complex production in insect cells by using polyproteins. *Methods Mol. Biol.* **1091**, 131–141 (2014).
14. Lee, Y. S. & Seong, B. L. Nucleotides in the panhandle structure of the influenza B virus virion RNA are involved in the specificity between influenza A and B viruses. *J. Gen. Virol.* **79**, 673–681 (1998).
15. Pflug, A., Guilligay, D., Reich, S. & Cusack, S. Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* <http://dx.doi.org/10.1038/nature14008> (this issue).
16. Flick, R., Neumann, G., Hoffmann, E., Neumeier, E. & Hobom, G. Promoter elements in the influenza vRNA terminal structure. *RNA* **2**, 1046–1057 (1996).
17. Pritlove, D. C., Poon, L. L., Devenish, L. J., Leahy, M. B. & Brownlee, G. G. A hairpin loop at the 5' end of influenza A virus virion RNA is required for synthesis of poly(A)⁺ mRNA *in vitro*. *J. Virol.* **73**, 2109–2114 (1999).
18. Newcomb, L. L. *et al.* Interaction of the influenza A virus nucleocapsid protein with the viral RNA polymerase potentiates unprimed viral RNA replication. *J. Virol.* **83**, 29–36 (2009).
19. Moeller, A., Kirchdoerfer, R. N., Potter, C. S., Carragher, B. & Wilson, I. A. Organization of the influenza virus replication machinery. *Science* **338**, 1631–1634 (2012).
20. Butcher, S. J., Grimes, J. M., Makeyev, E. V., Bamford, D. H. & Stuart, D. I. A mechanism for initiating RNA-dependent RNA polymerization. *Nature* **410**, 235–240 (2001).
21. Lescar, J. & Canard, B. RNA-dependent RNA polymerases from flaviviruses and Picornaviridae. *Curr. Opin. Struct. Biol.* **19**, 759–767 (2009).
22. Gong, P. & Peersen, O. B. Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proc. Natl Acad. Sci. USA* **107**, 22505–22510 (2010).
23. Deng, T., Vreede, F. T. & Brownlee, G. G. Different de novo initiation strategies are used by influenza virus RNA polymerase on its cRNA and viral RNA promoters during viral RNA replication. *J. Virol.* **80**, 2337–2348 (2006).
24. Mosley, R. T. *et al.* Structure of hepatitis C virus polymerase in complex with primer-template RNA. *J. Virol.* **86**, 6503–6511 (2012).
25. Plotch, S. J., Bouloy, M., Ulmanen, I. & Krug, R. M. A unique cap(m⁷GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell* **23**, 847–858 (1981).
26. Reguera, J., Weber, F. & Cusack, S. *Bunyaviridae* RNA polymerases (L-protein) have an N-terminal, influenza-like endonuclease domain, essential for viral cap-dependent transcription. *PLoS Pathog.* **6**, e1001101 (2010).
27. Sikora, D., Rocheleau, L., Brown, E. G. & Pelchat, M. Deep sequencing reveals the eight facets of the influenza A/HongKong/1/1968 (H3N2) virus cap-snatching process. *Sci. Rep.* **4**, 6181 (2014).
28. Datta, K., Wolkerstorfer, A., Szolar, O. H., Cusack, S. & Klumpp, K. Characterization of PA-N terminal domain of Influenza A polymerase reveals sequence specific RNA cleavage. *Nucleic Acids Res.* **41**, 349–353 (2013).
29. Hagen, M., Tiley, L., Chung, T. D. & Krystal, M. The role of template-primer interactions in cleavage and initiation by the influenza virus polymerase. *J. Gen. Virol.* **76**, 603–611 (1995).
30. Rao, P., Yuan, W. & Krug, R. M. Crucial role of CA cleavage sites in the cap-snatching mechanism for initiating viral mRNA synthesis. *EMBO J.* **22**, 1188–1198 (2003).
31. Geerts-Dimitriadou, C., Goldbach, R. & Kormelink, R. Preferential use of RNA leader sequences during influenza A transcription initiation *in vivo*. *Virology* **409**, 27–32 (2011).
32. Geerts-Dimitriadou, C., Zwart, M. P., Goldbach, R. & Kormelink, R. Base-pairing promotes leader selection to prime *in vitro* influenza genome transcription. *Virology* **409**, 17–26 (2011).
33. Braam, J., Ulmanen, I. & Krug, R. M. Molecular model of a eucaryotic transcription complex: functions and movements of influenza P proteins during capped RNA-primed transcription. *Cell* **34**, 609–618 (1983).
34. Li, M. L., Rao, P. & Krug, R. M. The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits. *EMBO J.* **20**, 2078–2086 (2001).
35. Honda, A., Mizumoto, K. & Ishihama, A. Two separate sequences of PB2 subunit constitute the RNA cap-binding site of influenza virus RNA polymerase. *Genes Cells* **4**, 475–485 (1999).
36. Shih, S. R. & Krug, R. M. Surprising function of the three influenza viral polymerase proteins: selective protection of viral mRNAs against the cap-snatching reaction catalyzed by the same polymerase proteins. *Virology* **226**, 430–435 (1996).
37. Lim, K. *et al.* Biophysical characterization of sites of host adaptive mutation in the influenza A virus RNA polymerase PB2 RNA-binding domain. *Int. J. Biochem. Cell Biol.* **53**, 237–245 (2014).
38. Ng, A. K. *et al.* Influenza polymerase activity correlates with the strength of interaction between nucleoprotein and PB2 through the host-specific residue K/E627. *PLoS ONE* **7**, e36415 (2012).
39. Coloma, R. *et al.* The structure of a biologically active influenza virus ribonucleoprotein complex. *PLoS Pathog.* **5**, e1000491 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the staff of the European Molecular Biology Laboratory (EMBL) eukaryotic expression and high-throughput crystallization facilities within the Partnership for Structural Biology (PSB) and members of the ESRF-EMBL Joint Structural Biology Group for help on European Synchrotron Radiation Facility (ESRF) beamlines. The work was supported by ERC Advanced Grant V-RNA (322586) and EU Grant FLU-PHARM (259751) to S.C. and partially by a Roche Postdoc Fellowship to S.R.

Author Contributions S.R., D.G. and T.L. did protein expression, purification, crystallization and activity assays. A.P. did crystallographic analysis. H.M. did electron microscopy and fitting to the mini-RNP electron microscopy map. M.N. calculated the first interpretable FluB polymerase electron density map. Using the polyprotein vector designed and provided by I.B., and with the help of D.H., S.C. designed the FluB polymerase construct. T.C., D.H., R.R. and S.C. have long-collaborated on studies of influenza polymerase. S.C. supervised the project, collected data, did crystallographic analysis and wrote the paper with input from S.R., D.G., A.P., H.M. and M.N.

Author Information Structure factors and co-ordinates have been deposited in the Protein Data Bank (PDB) under the accessions 4WSA (FluB form 1) and 4WRT (FluB form 2). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.C. (cusack@embl.fr).

METHODS

Construct. The influenza B/Memphis/13/03 polymerase heterotrimer was expressed as self-cleaving polypeptide (Extended Data Fig. 1). A codon-optimized synthetic construct (DNA2.0) with the composition $\text{GNH}_{\text{BstEII}} \text{GSGSENYFQ}_{\text{TEV}} \text{GSHH}_{\text{HHHHHH}}_{8\text{His-tag}} \text{GSGS-PA}$ (GenBank ID AAU94844) $\text{GSGSENYFQ}_{\text{TEV}} \text{GSGSGSGS-PB1}$ (GenBank ID AAU94857) $\text{GSGSENYFQ}_{\text{TEV}} \text{GSGSGSGS-PB2}$ (GenBank ID AAU94870) $\text{GWSHPQFEK}_{\text{strep-tag}} \text{GRSGR}_{\text{RsrII}}$ was cloned via BstEII and RsrII into the vector pKL-PBac¹³, which also contains coding sequences for tobacco etch virus (TEV) protease (5') and cyan fluorescent protein (CFP) (3'). (TEV cleavage site, His-tag and Strep-tag are underlined.)

Expression and purification. High Five insect cells expressing the target protein complex were resuspended in buffer A (50 mM Tris-HCl, 500 mM NaCl, 10% (v/v) glycerol and 5 mM BME, pH 8) supplemented with protease inhibitors (Roche, complete mini, EDTA-free), lysed by sonication and centrifuged at 30,000 r.p.m. for 30 min at 4 °C (rotor type 45 Ti, Beckman Coulter). Ammonium sulphate was added to the clarified supernatant (0.5 g ml⁻¹), the resulting precipitate collected by centrifugation as above and re-dissolved in buffer A supplemented with 20 mM imidazole. Soluble proteins were loaded on a nickel nitrilotriacetic acid (NTA) column (GE, FF crude) and bound proteins were eluted by 500 mM imidazole in buffer A. The target protein was loaded on a strep-tactin matrix (IBA, Superflow) and bound proteins eluted by 2.5 mM *d*-desthiobiotin in buffer A. Fractions containing the target protein were pooled and diluted with an equal volume of buffer B (50 mM HEPES/NaOH, 10% (v/v) glycerol and 2 mM Tris(2-carboxyethyl)phosphine (TCEP), pH 7.45) before loading on a heparin column (HiTrap Heparin HP, GE Healthcare). Proteins were eluted by a gradient of buffer B supplemented with 1 M NaCl, concentrated (Amicon Ultra, 50 kDa molecular mass cut-off) and further purified by size-exclusion chromatography (S200, GE Healthcare) in buffer C (50 mM HEPES/NaOH, 500 mM NaCl, 5% (v/v) glycerol and 2 mM TCEP, pH 7.45). Homogeneous monomeric FluB polymerase was concentrated as above and stored in aliquots at -80 °C. Protein concentration was determined by measuring the absorbance at 280 nm using the extinction coefficient 287,300 M⁻¹ cm⁻¹.

Crystallization. FluB polymerase was concentrated to 9 mg ml⁻¹ (37 μM) in a buffer containing 500 mM NaCl, 50 mM HEPES, pH 7.5, 5% glycerol and 2 mM TCEP, and mixed with 40 μM vRNA for crystallization in hanging drops at 4 °C. A trigonal crystal form (FluB1) was obtained by mixing polymerase with nucleotides 5–18 of the 3' end and 1–14 of the 5' end of the vRNA (IBA) in a condition containing 0.1 M bicine, pH 9.0, 10% MPD. Large (up to 150 μm) diamond-like crystals grew within a few days and diffracted to around 3.4 Å resolution but were very radiation-sensitive. The structure was solved with data at 6.5 Å resolution from a single heavy metal derivative obtained by soaking native crystals with 1 mM K₂PtCl₄ for 1 h. Selenomethionylated protein crystals were obtained in the same conditions as native ones. Polymerase with nucleotides 1–18 of the 3' end and 1–14 of the 5' end of the viral RNA gave thin hexagonal plates (form FluB2) in 1 M LiCl, 10% PEG 6000 and 0.1 M bicine, pH 9.0, that took 3–4 weeks to grow and diffracted to 2.7 Å resolution. All crystals were cryo-protected in mother liquor supplemented with 20% glycerol and flash-frozen in liquid nitrogen. Data was collected at 100 K on beamline ID23-1 at the European Synchrotron Radiation Facility (ESRF), equipped with a Pilatus 6M-F detector, at wavelengths of 0.9730 and 0.9792 Å for FluB1 and FluB2 crystals, respectively. All data were integrated and scaled with XDS⁴⁰.

Structure determination. A partial molecular replacement solution (LLG 334) was found with PHASER⁴¹ using the known PA-C-PB1-Nter (PDB codes 2ZN1 and 3CM8) and PB2 627 (PDB code 2VY7) domain structures initially both from FluA. The cap-binding and endonuclease domains could not be located even using the actual FluB domain structures (unpublished data). Nevertheless, ~22% of the complete structure was sufficient to identify around 20 platinum sites by inspection of a model-phased difference anomalous map. Several of the platinum peaks coincided with known positions of methionine residues. After scaling the platinum and native data sets, the platinum substructure was refined in SHARP⁴² to 7 Å and treated as SIRAS (single isomorphous replacement with anomalous), using the partial molecular replacement phases in the form of Hendrickson–Lattman coefficients. The final phasing statistics were phasing power (PP)_{anomalous} = 0.716, PP_{iso,centric} = 0.609, PP_{iso,acentric} = 0.714, figure of merit (FOM)_{centric} = 0.21, FOM_{acentric} = 0.36. Solvent flattening and phase extension to the full resolution of the native data (initially 3.7 Å and subsequently 3.4 Å) was then performed with SOLOMON⁴³ benefitting from the high solvent content of 73%. The resultant map had an overall correlation on |E|² of 81.4% and R_{factor} of 23.8%. The exceptionally good continuity of the map (Extended Data Fig. 3a–c) allowed immediate placing of known structures of the cap-binding, endonuclease and PB1–PB2 interface domains and revealed numerous additional secondary structures that could eventually be linked to trace almost the entire chain of each subunit as well as the vRNA. During model building and refinement with REFMAC⁴⁴ map sharpening with B_{factor} of -50 Å² was used to improve visibility of side chains. Accurate model building was aided by using four high resolution structures of FluB polymerase domains determined during the

course of this work (PA endonuclease at 2.1 Å resolution, PA-C/PB1-Nter at 2.4 Å resolution, PB2 cap-binding domain with m⁷GTP at 1.5 Å resolution, and the PB2 627-domain at 1.05 Å, unpublished data). Sequence assignment was verified by using the methionine positions located using the anomalous differences measured at 4.1 Å resolution from a seleno-methionylated polymerase crystal (Extended Data Fig. 3). The FluB2 crystal structure was determined by molecular replacement using the FluB1 structure. The C-terminal two-thirds of PB2 (PB2-C) is completely absent in the electron density map in this crystal form, although gel analysis shows PB2 to be intact and the crystal packing can accommodate PB2-C. When they became available, the higher resolution bat FluA and FluB2 structures enabled improvement in the quality of the FluB1 model with the help of secondary structure constraints derived using PROSMART⁴⁵. Full crystallographic statistics are given in Extended Data Table 1. Figures were drawn with Pymol⁴⁶. Ramachandran statistics, as calculated by Molprobity⁴⁷, are 93.7% (favoured), 0.6% (disallowed) for the FluB1 structure and 97.5% (favoured), 0.1% (disallowed) for the FluB2 structure.

Polymerase activity assays. A T7-transcribed 39-nucleotide mini-panhandle or equimolar mixture of separated synthetic 3' and 5' ends were used as vRNA (Extended Data Fig. 2), corresponding to the consensus promoter sequences for influenza B polymerase¹⁴.

For the ApG-primed replication assay, 0.5 μM protein, 0.5 μM vRNA, 0.5 mM ApG, 0.4 mM GTP/CTP, 1 mM ATP, 0.04 mM UTP, ³²P-UTP and 0.8 U μl⁻¹ Ribolock, in buffer (150 mM NaCl, 50 mM HEPES, pH 7.5, 5 mM MgCl₂ and 2 mM TCEP) were mixed and incubated at 30 °C for 2 h.

For the cap-dependent transcription assay, 0.5 μM protein, 0.5 μM vRNA, 0.4 mM GTP/CTP/UTP, 1 mM ATP, ³²P-labelled capped RNA in the same buffer (150 mM NaCl, 50 mM HEPES, pH 7.5, 5 mM MgCl₂ and 2 mM TCEP) were mixed and incubated at 30 °C for 2 h. For this purpose, a 5' diphosphate synthetic 20-base RNA, 5'-ppAAUCUAUAAUAGCAUUAUCC-3' (Chemgenes), was capped by incubating with vaccinia virus capping enzyme (purified in house following ref. 48) and 20 μM SAM, ³²P-GTP, 50 mM Tris, pH 8.0, 6 mM KCl, 1.25 mM MgCl₂ and 0.8 U μl⁻¹ Ribolock.

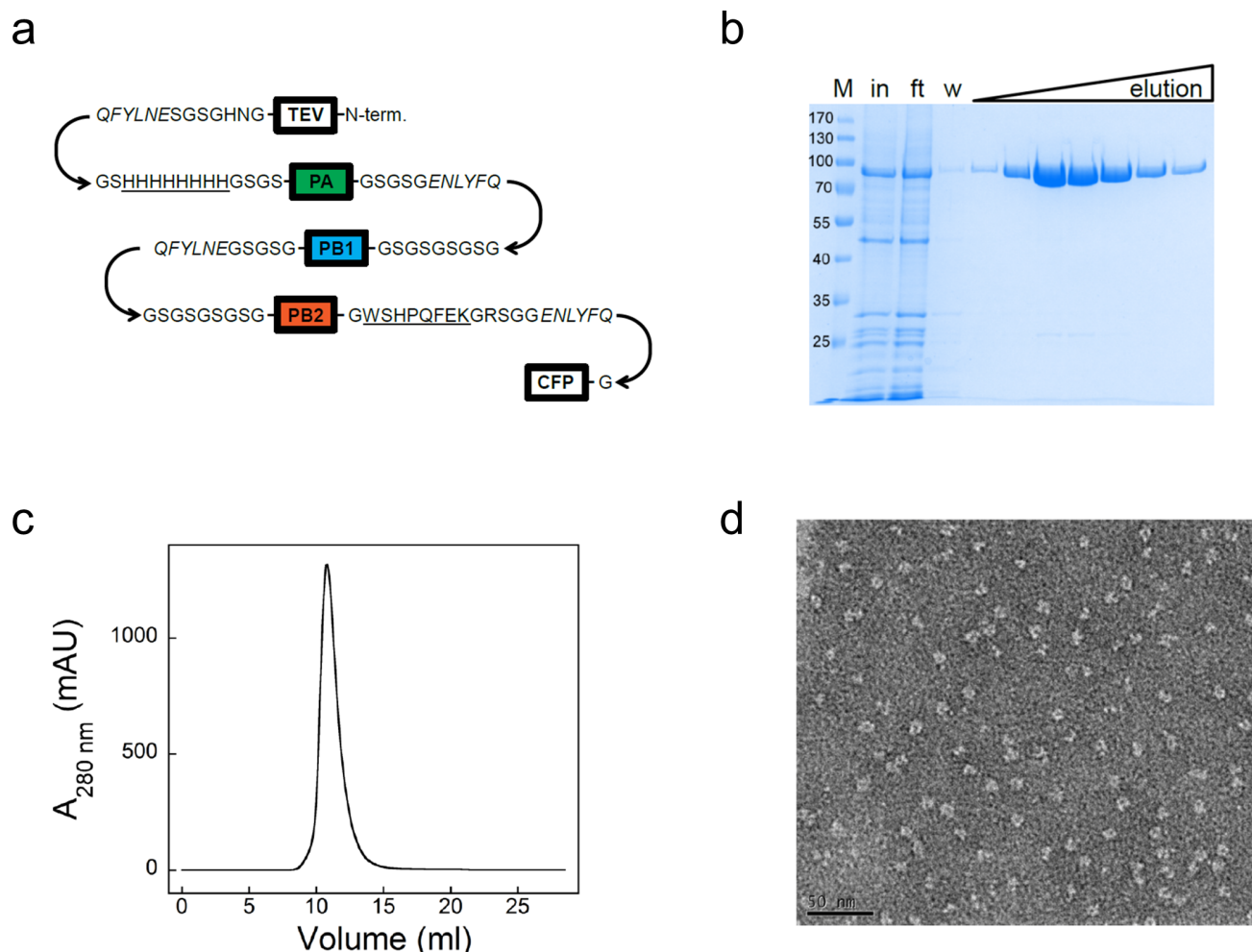
For the endonuclease assay, transcription mix lacking NTPs was incubated at 30 °C for 2 h. Samples were separated on 7 M urea, 20% acrylamide gel in TBE buffer, exposed on a Storage Phosphor screen and read with a Typhoon scanner.

For the time course of unprimed and ApG-primed vRNA replication, 0.5 μM FluB polymerase was mixed with 1 μM 39-nucleotide vRNA mini-panhandle template, NTPs (1 mM ATP, 0.4 mM GTP, 0.4 mM CTP and 0.04 mM UTP) and 0.12 μCi μl⁻¹ ³²P-UTP, in the absence or presence of 0.5 mM ApG. Reactions were incubated at 30 °C and samples were analysed on a 20% acrylamide, 7 M urea denaturing gel after 0, 2, 5, 10, 15, 20, 30, 40 and 50 min, 1, 2 and 3 h.

Fitting to the mini-RNP electron microscopy map. Influenza polymerase (FluB2 model) and nine influenza A nucleoproteins (PDB code 2IQH)⁴⁹ were rigidly fitted into the 18 Å mini-RNP cryo-EM reconstruction³⁹ using chimera fit-in-map module⁵⁰ and VEDA⁵¹. Map scaling was optimized by the cross-correlation between the model and map for different pixel sizes as implemented in VEDA. Down-scaling the electron microscopy map from 2.8 Å/pixel to 2.4 Å/pixel improved the cross-correlation and fit quality considerably. The fitting of the nine nucleoproteins follow the model previously proposed³⁹, with each nucleoprotein and loop 402–428 of its neighbour being considered as a rigid entity to maintain the nucleoprotein–nucleoprotein interaction mode. For polymerase fitting, different starting positions of the PA–PB1 heterodimer with only 1–32 of PB2 (FluB2 model) were used for rigid body fitting using the chimera fit-in-map module and allowed to identify one preferred rigid fit position. Finally, the model was refined with a simultaneous rigid fit of the polymerase and the nine nucleoproteins using VEDA.

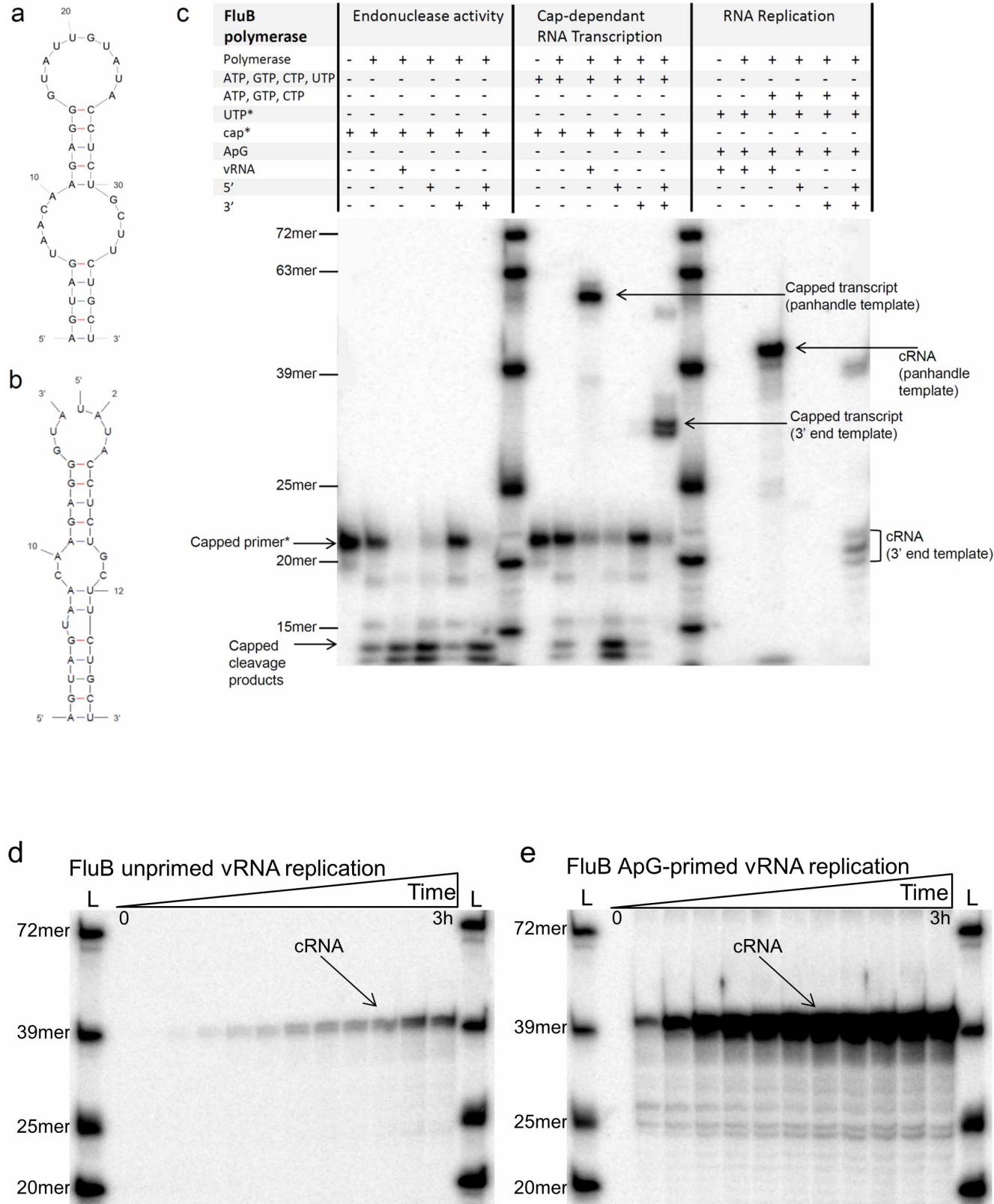
- Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D* **66**, 133–144 (2010).
- Read, R. J. Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr. D* **57**, 1373–1382 (2001).
- Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D* **59**, 2023–2030 (2003).
- Abrahams, J. P. & Leslie, A. G. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D* **52**, 30–42 (1996).
- Murshudov, G. N. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
- Nicholls, R. A., Fischer, M., McNicholas, S. & Murshudov, G. N. Conformation-independent structural comparison of macromolecules with ProSMART. *Acta Crystallogr. D* **70**, 2487–2499 (2014).
- DeLano, W. L. The PyMOL Molecular Graphics System; <http://www.pymol.sourceforge.net> (Schrödinger, LLC, 2002).
- Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
- De la Peña, M., Kyrieles, O. J. & Cusack, S. Structural insights into the mechanism and evolution of the vaccinia virus mRNA cap N7 methyl-transferase. *EMBO J.* **26**, 4913–4925 (2007).

49. Ye, Q., Krug, R. M. & Tao, Y. J. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* **444**, 1078–1082 (2006).
50. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comp. Chem.* **25**, 1605–1612 (2004).
51. Siebert, X. & Navaza, J. UROX 2.0: an interactive tool for fitting atomic models into electron-microscopy reconstructions. *Acta Crystallogr. D* **65**, 651–658 (2009).
52. Esnouf, R. M. Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. *Acta Crystallogr. D* **55**, 938–940 (1999).
53. Area, E. *et al.* 3D structure of the influenza virus polymerase complex: localization of subunit domains. *Proc. Natl Acad. Sci. USA* **101**, 308–313 (2004).
54. Wakai, C., Iwama, M., Mizumoto, K. & Nagata, K. Recognition of cap structure by influenza B virus RNA polymerase is less dependent on the methyl residue than recognition by influenza A virus polymerase. *J. Virol.* **85**, 7504–7512 (2011).



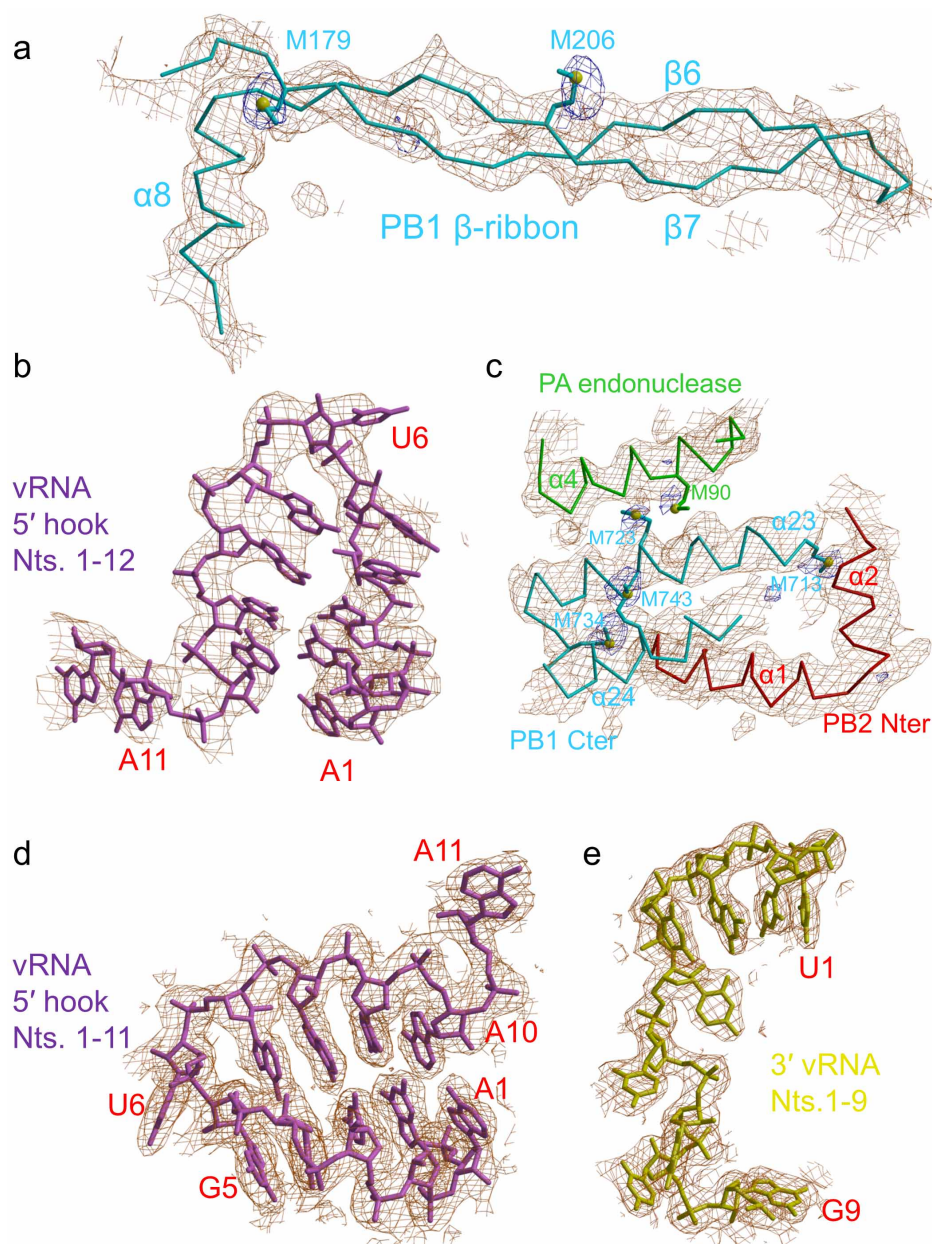
Extended Data Figure 1 | Production and characterization of influenza B polymerase heterotrimer. **a**, Schematic of the self-cleaving polyprotein construct used to express recombinant influenza B heterotrimeric polymerase in insect cells. N-terminally it encodes the tobacco etch virus (TEV) protease that cleaves C-terminal to the amino-acid sequence ENLYFQ (in *italics*) and releases N-terminally His-tagged PA, PB1, C-terminally Strep-tagged PB2 and cyan fluorescent protein (CFP) for facilitated expression monitoring. Arrows indicate the N-to-C-terminal direction and the termini of each mature protein. The histidine and streptavidin tags are underlined. **b**, Using the PB2 C-terminal strep-tag, most contaminating proteins could be separated from the polymerase as judged by 10% SDS-PAGE followed by Coomassie blue staining. Lanes 'M' contain the protein markers (molecular masses indicated); 'in', 'ft' and 'w' denote the input, flow-through and wash of the engineered

streptavidin (strep-tactin) column, respectively, and 'elution' indicates the re-mobilization of bound heterotrimeric polymerase by a sharp gradient of *d*-desthiobiotin. The three subunits, PA (85.7 kDa), PB1 (86.1 kDa) and PB2 (90.8 kDa), run together on the gel. **c**, After ammonium sulphate precipitation, IMAC, strep-tactin affinity and heparin chromatography, the final purification step consists of size-exclusion chromatography. The elution profile (monitored by the absorbance at 280 nm) with a single and nearly symmetric peak suggests a homogeneous and monomeric polymerase complex. **d**, Recombinant influenza B polymerase was analysed by electron microscopy following negative staining with sodium silico-tungstate of 0.02 mg ml^{-1} protein sample. The image demonstrates that the sample is homogeneous and monodisperse with a V- or doughnut-like shape with a central cavity.



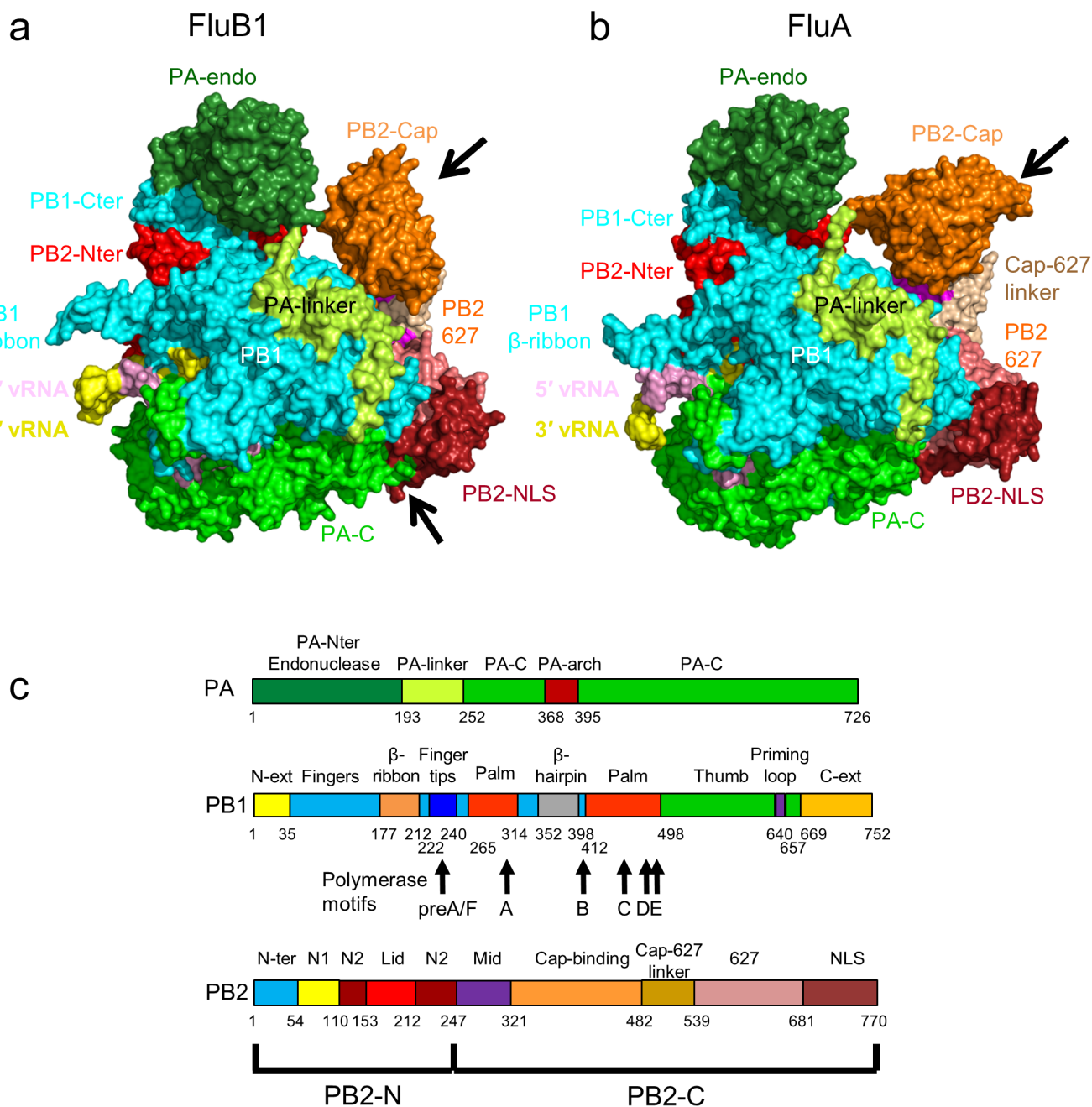
Extended Data Figure 2 | Endonuclease, transcription and replication activities of FluB polymerase. **a**, Schematic of mini-panhandle vRNA: 5'-pppAGUAGUAACAAGAGGGUAUUGUAUACCUCUGCUUCUGCU-3'. **b**, Schematic of separate 5' and 3' ends: 5': 5'-pAGUAGUAACAAGA GGGUA-3'; 3': 5'-UAUACCUCUGCUUCUGCU-3'. **c**, Endonuclease, cap-dependent transcription and ApG-primed replication assays. Cleavage of the cap donor is visible in lanes 2–6 and enhanced in the presence of the 5' end, but not the 3' end. Capped transcripts are visible in lanes 10 (from vRNA

panhandle template) and 13 (from separated 5' and 3' vRNA ends) as well as cRNA produced in lanes 17 and 20. Markers, with size shown on the left, are RNA ladders labelled with ³²P-pCp nucleotide. **d**, **e**, Time course of unprimed (**d**) and ApG-primed (**e**) vRNA replication by influenza B polymerase. The products of replication (cRNA) are indicated with an arrow. Ladders (lanes L) are ³²P-pCp nucleotide-labelled RNA oligomers. ApG-primed replication is more efficient than unprimed replication.



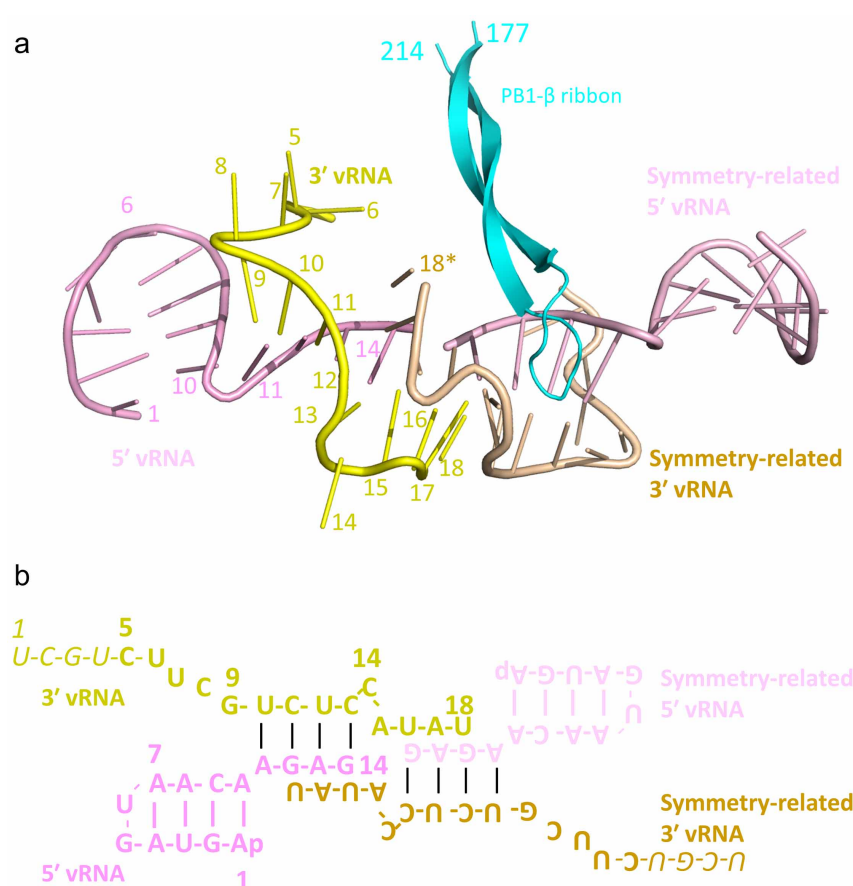
Extended Data Figure 3 | Examples of electron density map for FluB polymerase. **a–c**, Initial platinum SIRAS-phased and phase-extended experimental map at 3.6 Å resolution contoured at 1.1 σ (brown) with superposed final model for the FluB1 crystal form. Also shown is the final model-phased selenium anomalous difference map at 4.1 Å resolution

contoured at 3.2 σ (purple) highlighting methionine positions. **a**, PB1 β -ribbon. **b**, vRNA 5' hook. **c**, PA–PB1–PB2 helical interface. **d**, **e**, final $2F_o - F_c$ omit map at 2.7 Å resolution for the FluB2 crystal form contoured at 1.1 σ . **d**, vRNA 5' hook nucleotides 1–11. **e**, vRNA 3' end nucleotides 1–9. Figures drawn with Bobscrip⁵².



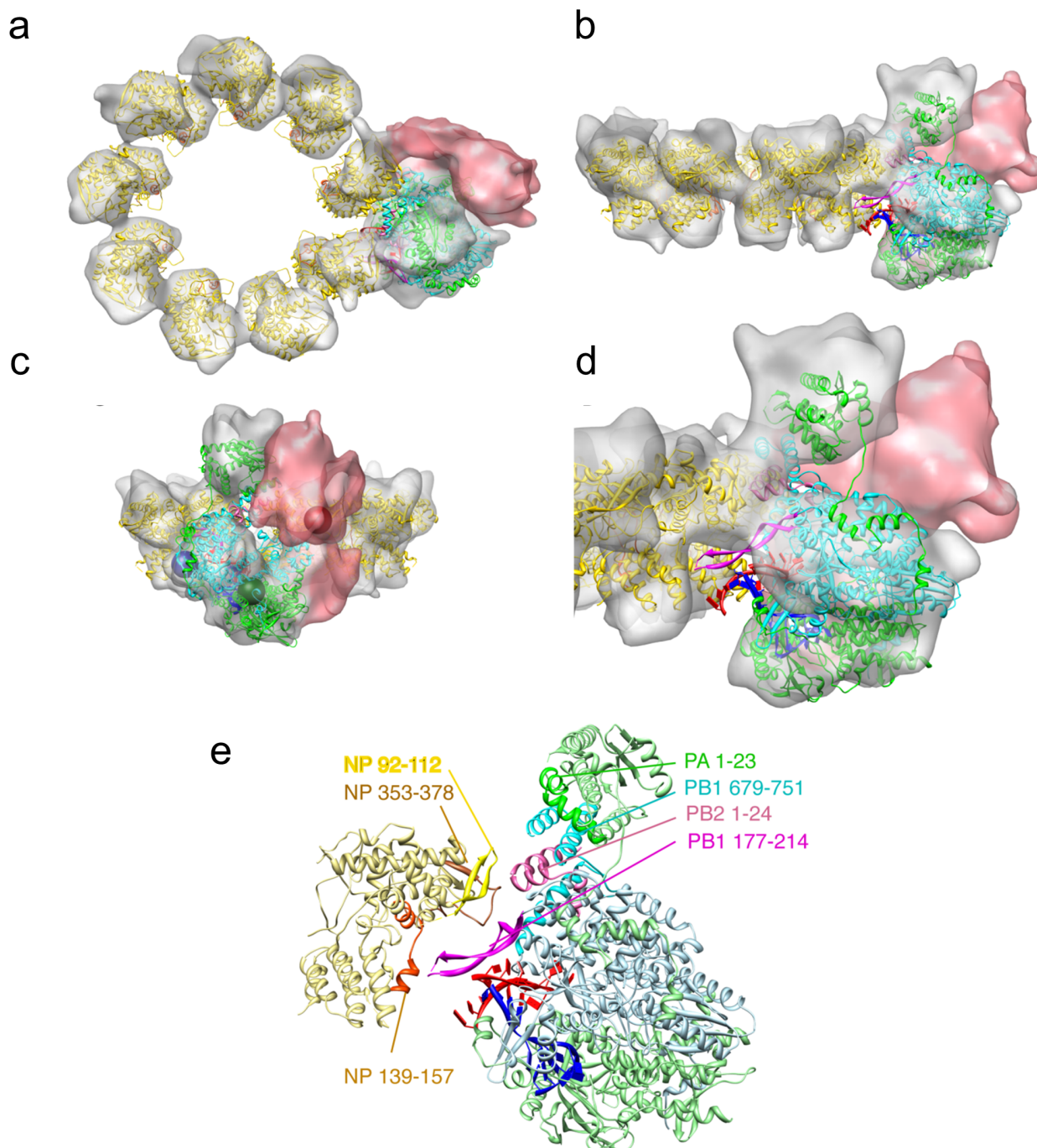
Extended Data Figure 4 | Comparison of FluB and bat FluA polymerase structures. **a**, Surface diagram of FluB1 structure coloured as in **c** except that PA-C, PB1 and PB2-N are uniformly green, cyan and red, respectively. The bottom black arrow indicates the extra 12 C-terminal residues of FluB PA that extend the PA C-terminal helix compared to FluA, so that it directly contacts the PB2-NLS domain that is consequently orientated slightly differently from in FluA polymerase. **b**, As in **a** but for bat FluA structure. Arrows highlight the 70° difference in orientation of the cap-binding domain. The structural similarity between FluA and FluB polymerases (LSQMAN, cut-off 3.5 Å) is as follows. PA: 630 Cα atoms aligned, of which 38.6% are identical with root mean squared deviation (r.m.s.d.) 1.34 Å; PB1: 703 Cα atoms

aligned, of which 61.3% are identical with r.m.s.d. 1.06 Å; PB2: 428 Cα atoms aligned, of which 40.6% are identical with r.m.s.d. 1.46 Å (excluding the cap-binding domain), and, taking into account the cap-binding domain rotation, 622 Cα atoms aligned, of which 39.0% are identical with r.m.s.d. 1.54 Å. **c**, Subunit domain structure of influenza B polymerase with names and extended colour scheme, showing the positions of the PB1 polymerase motifs. Note that for PB1, the FluB numbering compared to FluA is the same from 1–399 and is thereafter +1. For PB2, FluB is +2 from 1–469 and +1 from 470–628. For PA it is more complicated owing to several short insertions and deletions. See Supplementary Fig. 1.



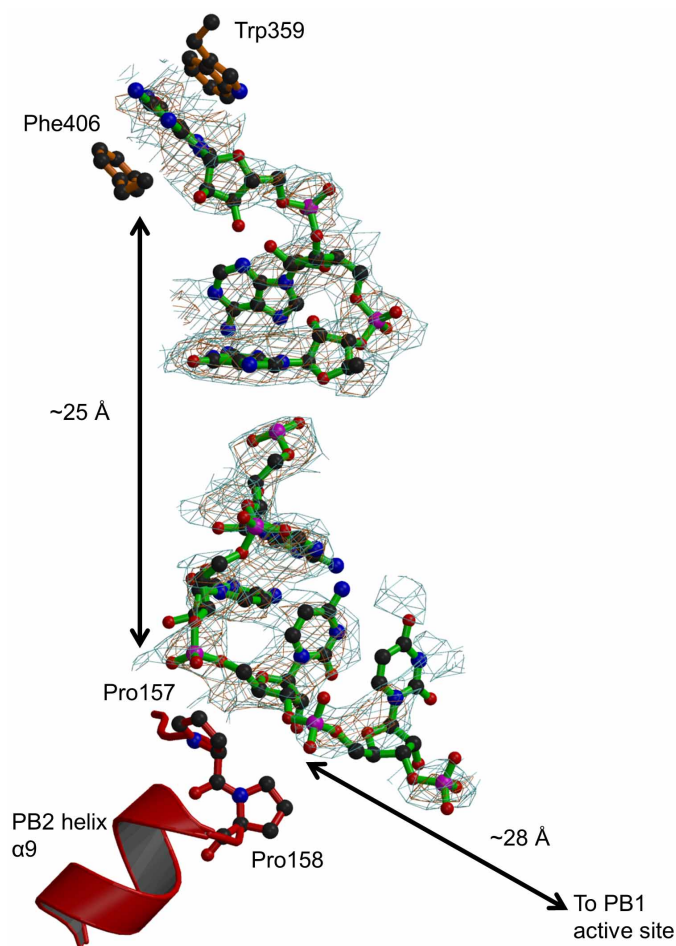
Extended Data Figure 5 | RNA-RNA crystal contact in FluB1 crystal form.
a, Cartoon of 5' and 3' vRNA ends (left, pink and yellow, respectively) interacting with crystallographic two-fold symmetry-related vRNA (right, pale pink and wheat, respectively). The PB1 β-ribbon (cyan) of the left-hand

polymerase molecule interacts with the symmetry-related vRNA. **b**, Simplified diagram showing vRNA sequence and secondary structure in the FluB1 crystal form including vRNA-mediated crystal contact.



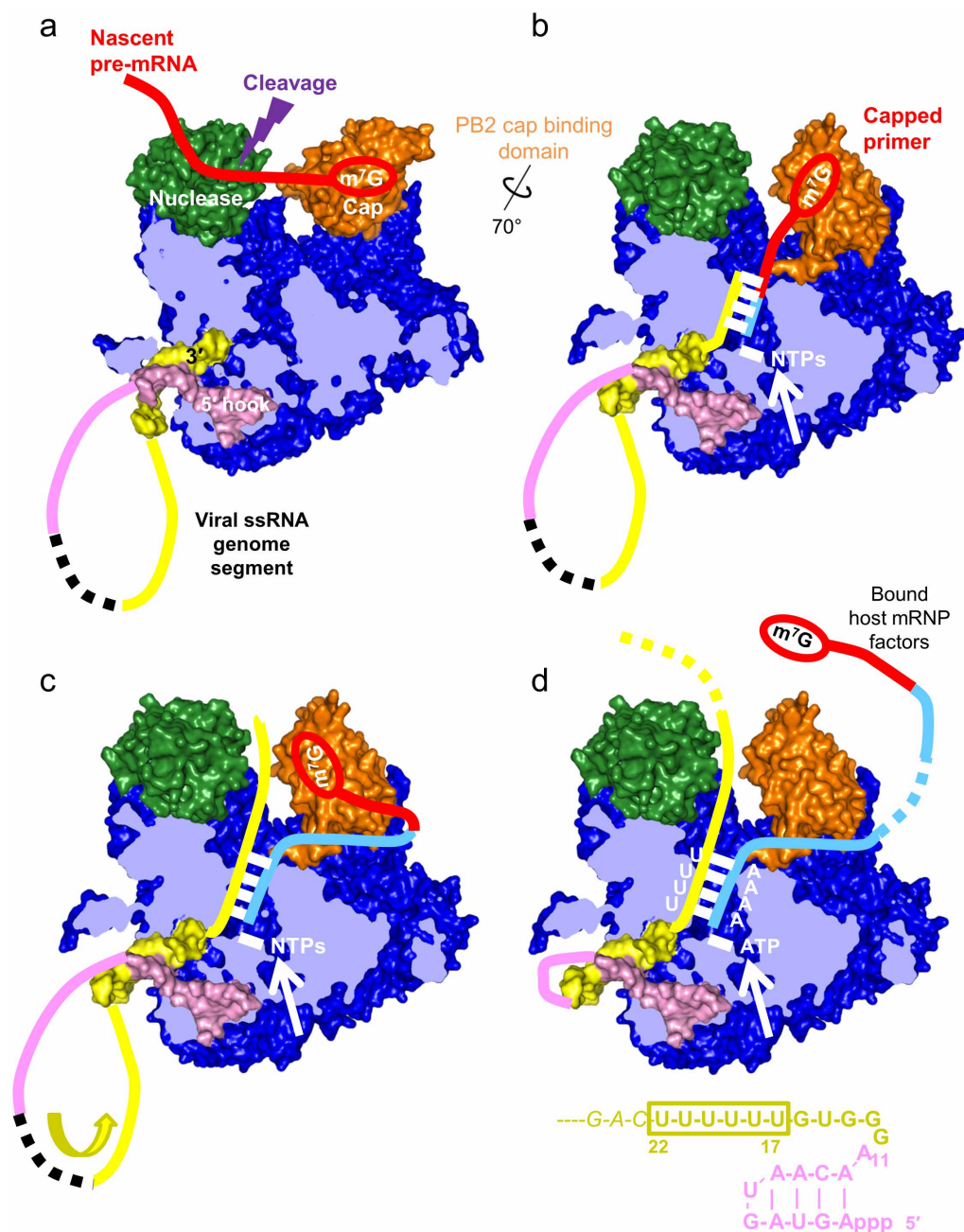
Extended Data Figure 6 | Polymerase fitting into the mini-RNP electron microscopy map. **a, b,** Top (**a**) and side (**b**) view of influenza A mini-RNP pseudo-atomic model with rescaled electron density³⁹. PA, PB1 and PB2 (1–32 only) are shown as ribbons and coloured in green, cyan and rose, respectively. Unfilled electron density, likely to contain the rest of PB2, is shown in transparent rose. Nucleoproteins are shown in yellow ribbons, with the nucleoprotein–nucleoprotein interacting loop (residues 402–428) in orange. The vRNA 5' and 3' ends are shown in dark blue and red, respectively. **c,** Front view of influenza A mini-RNP pseudo-atomic model. The positions of

antibody and tag labelling corresponding to domains of PA, PB1 and PB2 are shown as dark green, dark blue and dark rose spheres, respectively, as localized previously⁵³. **d,** Close-up view of **b**. The PB1 β -ribbon (residues 177–214, purple) is located close to one of the proximal nucleoproteins and the vRNA. **e,** Putative interactions between the proximal nucleoprotein and polymerase. Nucleoprotein elements proposed for polymerase interaction are indicated in yellow, brown and orange. Polymerase interacting elements are shown in green, cyan, rose and magenta.



Extended Data Figure 7 | Residual electron density in the FluB1 crystal form mimicking capped primer binding to the PB2 cap-binding domain.

Residual $m2F_o - F_c$ (blue mesh at 0.9σ) and $mF_o - F_c$ (orange mesh at 2.5σ) electron density showing RNA-like density bound in the cap-binding site in the FluB1 crystal form. The low resolution and partial occupancy do not allow identification of the RNA and the discontinuous model shown is for illustrative purposes only. Owing to the rigorous purification procedure it is unlikely to be insect cell-capped RNA that is trapped on the polymerase. More likely it derives from the input vRNA used in crystallization, possibly partially digested by the endonuclease that generates 3' ends. That this RNA could even be uncapped is explicable by the fact that the FluB cap-binding domain, unlike that of FluA, promiscuously binds both methylated and unmethylated guanosine⁵⁴. Indeed, the density seems to be better fit with a free 3' end sandwiched between Phe 406 and Trp 359 in the cap-binding site rather than a capped 5' end. As the primer emerges from the cap-binding site it is initially channelled on one side by the base of the 424-loop, and on the other by residues 518–522 of the cap-627 linker. Further down, the extended 424-loop continues to guide the RNA, as well as, on the other side, the projecting N-terminal end of PB2 helix $\alpha 9$ (155-EMPPDE in FluB), with the double proline forcing the RNA into a $\sim 90^\circ$ bend. Arg 425 and Arg 438 are well placed to interact with phosphates and one base seems to stack on the Glu 155–Arg 217 salt bridge. Conserved basic residues on PB2 N2 domain strands $\beta 7$, 144-Arg-Lys-Arg (FluA 142-Arg-Lys-Arg), and $\beta 8$, 216-Arg-Arg-Arg-Phe (FluA 214-Arg-Thr-Arg-Phe), are also likely to be involved. Straight-line distances from the cap-binding site to the bend and from the bend to the PB1 active site are indicated. See also Fig. 5.



Extended Data Figure 8 | Schematic diagram of steps in cap-dependent transcription by influenza virus polymerase. a, Cap-snatching from host pre-mRNA (red). The m⁷G cap is bound by the cap-binding domain (orange, orientated as in the FluA structure) and the pre-mRNA cleaved 10–14 nucleotides downstream by the endonuclease (green). The single-stranded vRNA genome is bound by its 5' (hook, pink) and 3' (template, yellow) ends to the polymerase (blue, depicted as a cutaway section). **b,** Transcription initiation. The cap-binding domain rotates to the position observed in the FluB1 structure directing the capped primer into the PB1 active site, where it potentially makes limited base pairs with the extremity of the template. Template-directed NTP addition (white) extends the host sequences (red) with

virally encoded sequences (cyan). Note that in **b–d** additional conformational changes in the polymerase are expected, but not depicted since they are currently unknown. **c,** Transcription elongation. Transcription elongation proceeds, eventually leading to the release of the cap from the cap-binding domain (**d**) and the binding of host mRNP factors. **d,** Polyadenylation by stuttering. After most of the vRNA template has been translocated through the polymerase, only a tight turn connects it to the bound 5'-hook. The nucleotide sequence of this region is given at the bottom. This places the 5' proximal oligo-U stretch in the PB1 active site allowing poly(A) tail synthesis by a stuttering mechanism in which the template is no longer translocated but the product strand is able to slip.

Extended Data Table 1 | Data collection and refinement statistics for FluB polymerase structures

	FluB1 3' end: 5-18 5' end: 1-14	FluB2 3' end: 1-18 5' end: 1-14
Data collection		
Space group	$P3_221$	$P6_222$
Cell dimensions		
a, b, c (Å)	199.70, 199.70, 252.68	207.37, 207.37, 345.69
α, β, γ (°)	90.0, 90.0, 120.0	90.0, 90.0, 120.0
Number of crystals	4	1
Resolution (Å)	50.0-3.40(3.52-3.40)*	50.0-2.70(2.80-2.70)
R_{meas}	12.1 (130.1)	25.2 (142.9)
$I/\sigma I$	12.6 (1.5)	10.3 (1.9)
Completeness (%)	98.5 (89.9)	99.9 (100.0)
Redundancy	11.8 (5.8)	10.1 (10.3)
Refinement		
Resolution (Å)	50.0-3.40(3.49-3.40)	50.0-2.70(2.77-2.70)
No. reflections	79266	120050
$R_{\text{work}}/R_{\text{free}}$	22.9/26.5 (40.2/42.7)	17.3/21.1 (32.3/34.6)
No. atoms		
Protein	17351	13548
RNA	589	615
Water	-	809
B-factors (Å ²)		
Protein	158.5	43.5
RNA	159.6	43.5
Water	126.1	52.7
R.m.s deviations		
Bond lengths (Å)	0.003	0.003
Bond angles (°)	0.657	0.683

*Highest resolution shell is shown in parenthesis.

Extended Data Table 2 | Direct polar polymerase–vRNA contacts for the FluB2 structure

Calculated with CONTACT (CCP4i) with 3.5 Å cutoff.

PA (chain B) - 5' vRNA (chain V)

Residue	Atom		RNA base	Atom	Distance (Å)
Lys 330B	NZ	...	A	1V OP1	3.21 ***
Gly 367B	N	...	A	1V N1	3.44 *
Gly 367B	O	...	A	10V O2'	2.75 ***
Gly 369B	N	...	A	11V OP2	2.76 ***
Leu 370B	N	...	A	11V OP1	3.23 ***
Thr 371B	N	...	A	11V OP1	3.02 ***
Thr 371B	OG1	...	A	10V OP2	2.91 ***
		...	A	10V N7	3.45 *
Gln 504B	NE2	...	A	11V O4'	3.33 *
		...	A	11V O5'	3.05 ***
His 506B	NE2	...	A	11V N7	3.32 *
Val 513B	N	...	C	9V O2'	2.89 ***
Thr 515B	OG1	...	A	1V N1	2.88 ***
Arg 558B	NH2	...	U	3V OP1	2.98 ***
Val 559B	O	...	G	2V O3'	3.41 *
Asn 560B	O	...	G	2V O2'	3.21 ***
Gly 561B	O	...	U	3V O3'	3.41 *
		...	U	3V O2'	3.42 *
Gln 566B	NE2	...	A	4V OP1	2.94 ***
Asn 692B	ND2	...	G	5V O6	2.90 ***

PB1 (chain P) - 5' vRNA (chain V)

His 32P	NE2	...	G	5V OP1	3.10 ***
His 32P	O	...	A	7V O4'	3.43 *
Thr 34P	N	...	A	8V OP1	3.20 ***
Tyr 38P	OH	...	U	6V O5'	3.32 *
Lys 365P	NZ	...	C	9V OP2	2.92 ***

PA (chain B) - 3' vRNA (chain T, numbered from 3' end)

His 506B	ND1	...	G	9T O6	3.01 ***
Arg 508B	NE	...	U	7T OP1	3.02 ***
Arg 508B	NH1	...	U	10T O2	3.45 *
Arg 508B	NH2	...	U	10T O2'	2.49 ***
		...	U	10T O2	3.23 ***
Lys 564B	NZ	...	C	8T OP2	2.79 ***

PB1 (chain P) - 3' vRNA (chain T, numbered from 3' end)

Gln 127P	OE1	...	U	6T N3	2.97 ***
		...	U	6T O2	3.33 *
Val 133P	O	...	U	4T N3	3.42 *
		...	U	4T O4	3.07 ***
Arg 135P	NE	...	U	4T O4	3.01 ***
Arg 135P	NH2	...	U	4T O4	3.16 ***
		...	G	3T O6	3.15 ***
Asn 136P	O	...	U	6T N3	2.89 ***
Val 184P	O	...	G	3T N2	2.83 ***
Asn 186P	N	...	C	2T O2	2.71 ***
Asn 186P	ND2	...	C	2T N3	3.26 ***
Arg 203P	NH2	...	U	1T OP2	2.41 ***
Arg 350P	NH2	...	U	4T O4'	2.90 ***
Asn 670P	OD1	...	U	7T OP2	3.35 *
Asn 670P	ND2	...	U	7T OP2	2.92 ***
Arg 671P	N	...	G	9T OP1	2.88 ***
Arg 671P	NH1	...	U	10T OP1	2.67 ***
Ser 672P	N	...	U	10T O3'	3.39 *
Ser 672P	OG	...	U	10T O2'	2.82 ***
		...	U	7T OP2	2.61 ***
Asn 675P	ND1	...	C	11T O2'	2.54 ***
		...	C	11T O2	3.17 ***

PB2 (chain E) - 3' vRNA (chain T, numbered from 3' end)

Thr 38E	O	...	U	7T N3	2.82 ***
		...	U	7T O2	3.27 ***
Arg 40E	N	...	U	7T O4	3.25 ***
Arg 40E	NH1	...	U	7T O4'	2.82 ***
Arg 40E	NH2	...	U	6T O2	3.45 *
Arg 40E	O	...	C	8T N4	2.99 ***
Glu 42E	OE2	...	C	8T N3	3.47 *
Arg 48E	NH1	...	C	8T OP2	2.71 ***
Arg 48E	NH2	...	C	8T O4'	3.26 ***
Trp 51E	NE1	...	G	9T O2'	2.74 ***

The exclusion of a significant range of ages in a massive star cluster

Chengyuan Li^{1,2,3}, Richard de Grijs^{1,2} & Licai Deng³

Stars spend most of their lifetimes on the main sequence in the Hertzsprung–Russell diagram. The extended main-sequence turn-off regions—containing stars leaving the main sequence after having spent all of the hydrogen in their cores—found in massive (more than a few tens of thousands of solar masses), intermediate-age (about one to three billion years old) star clusters^{1–8} are usually interpreted as evidence of internal age spreads of more than 300 million years^{2,4,5}, although young clusters are thought to quickly lose any remaining star-forming fuel following a period of rapid gas expulsion on timescales of order 10^7 years^{9,10}. Here we report, on the basis of a combination of high-resolution imaging observations and theoretical modelling, that the stars beyond the main sequence in the two-billion-year-old cluster NGC 1651, characterized by a mass of about 1.7×10^5 solar masses³, can be explained only by a single-age stellar population, even though the cluster has a clearly extended main-sequence turn-off region. The most plausible explanation for the existence of such extended regions invokes a population of rapidly rotating stars, although the secondary effects of the prolonged stellar lifetimes associated with such a stellar population mixture are as yet poorly understood. From preliminary analysis of previously obtained data, we find that similar morphologies are apparent in the Hertzsprung–Russell diagrams of at least five additional intermediate-age star clusters^{2,3,5,11}, suggesting that an extended main-sequence turn-off region does not necessarily imply the presence of a significant internal age dispersion.

We obtained archival Hubble Space Telescope Wide Field Camera 3 observations of the NGC 1651 field in the F475W ('B') and F814W ('I') broadband filters (Methods). The corresponding colour–magnitude diagram, that is, the observational counterpart of the Hertzsprung–Russell diagram, is shown in Fig. 1. When stars have exhausted their core hydrogen supply, hydrogen fusion continues in a shell outside the

stellar core. At this stage, stars leave the main sequence and evolve onto the subgiant branch. The colour–magnitude diagram of NGC 1651 exhibits a clearly extended main-sequence turn-off (simply referred to as a 'turn-off' in what follows) and a very narrow subgiant branch. This is surprising, given the corresponding, far-reaching implications of our interpretation of such extended turn-offs in the context of star cluster evolution.

Star clusters more massive than a few tens of thousands of solar masses were, until recently, considered single-generation ('simple') stellar populations. It was thought that all of their member stars had formed approximately simultaneously from molecular gas originally confined to a small volume of space. As a consequence, all cluster stars would thus have similar ages, a very narrow range in chemical composition and individual stellar masses that followed the initial mass function, that is, the stellar mass distribution at the time of star birth. In the past decade, however, a consensus has emerged that massive star clusters are not ideal simple stellar populations^{12–18}. Deviations from the simple-stellar-population model in resolved star clusters are most readily discerned by reference to their colour–magnitude diagrams, and in particular to their turn-off regions.

Taking a simplistic, direct approach, we obtain best fits to the blue and red edges of the extended turn-off by matching the best set of theoretical stellar isochrones¹⁹ available at present to the observed stellar distribution. The best-fitting isochrones bracketing the data range from $\log[t(\text{yr})] = 9.24$ to $\log[t(\text{yr})] = 9.34$ (where t represents the stellar population's age), for a stellar metal (iron) abundance of $[\text{Fe}/\text{H}] = -0.52$ dex (ref. 20), a reddening of $E(B - V) = 0.11$ mag and a distance modulus of $(m - M)_0 = 18.46$ mag (ref. 21). Figure 1 shows the 'cleaned' colour–magnitude diagram (Methods). The lines represent the best-fitting theoretical isochrones covering the cluster's extended turn-off region. Although this region is well described by adoption of

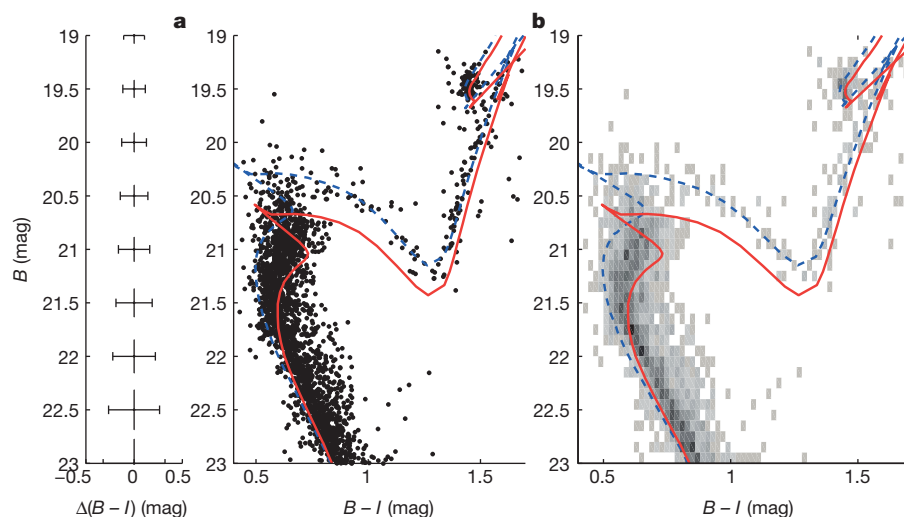


Figure 1 | NGC 1651's stellar distribution in colour-magnitude space. **a**, Colour-magnitude diagram, including typical 3σ photometric uncertainties. The blue dashed and red solid lines represent isochrones for $\log[t(\text{yr})] = 9.24$ and $\log[t(\text{yr})] = 9.34$, respectively. **b**, Corresponding number density ('Hess') diagram.

¹Kavli Institute for Astronomy and Astrophysics, Peking University, Yi He Yuan Lu 5, Hai Dian District, Beijing 100871, China. ²Department of Astronomy, Peking University, Yi He Yuan Lu 5, Hai Dian District, Beijing 100871, China. ³Key Laboratory for Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Chaoyang District, Beijing 100012, China.

an age dispersion of approximately 450 Myr, the cluster's subgiant-branch stars are predominantly confined to the youngest isochrone. The 15 subgiant-branch stars in the NGC 1651 core region (with a radius of ≤ 20 arcsec ≈ 5 pc; see Methods) are confined to an even narrower distribution along the subgiant branch than is the full sample of 38 stars selected using the box in Fig. 2a, c. This indicates that the narrow width of the subgiant branch does not depend on position in the cluster. However, a 450 Myr age spread would also require a significant broadening of the cluster's subgiant branch. This is why our discovery of a subgiant branch in NGC 1651 with a very narrow stellar distribution is surprising, which thus immediately leaves us with a conundrum.

To assess the association of our subgiant-branch stars with either the youngest or the oldest isochrone, we first adopt the $\log[t(\text{yr})] = 9.24$ isochrone as our baseline and calculate the individual deviations, ΔB (mag), for all subgiant-branch stars. We subsequently adopt the $\log[t(\text{yr})] = 9.34$ isochrone as our fiducial locus. The blue and orange regions in Fig. 2b, d correspond to the typical deviations expected for subgiant-branch stars associated with the youngest and oldest isochrones, respectively, assuming a 3σ magnitude dispersion of $\Delta B = 0.12$ mag. Thirty of the 38 stars (14 of 15 stars in the core) are associated with the youngest isochrone. Only a single subgiant-branch star, located outside the cluster's core region, might statistically be associated with the region in parameter space defined by the oldest isochrone. If we directly use the observed

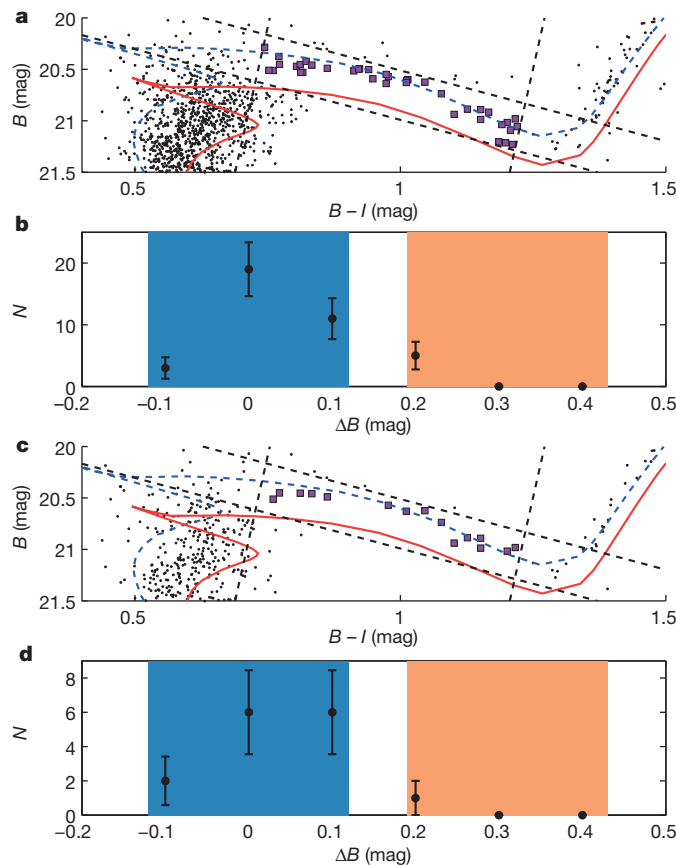


Figure 2 | Comparison of the observed stellar distribution with the expectations of a 450 Myr spread in cluster internal age. **a**, Region of the colour-magnitude diagram covering the extended turn-off and the subgiant branch (indicated by the black dashed lines; purple squares, subgiant-branch stars). The blue dashed and red solid isochrones are as in Fig. 1. **b**, Number distribution, N (including 1σ standard deviations), of the deviations in magnitude, ΔB , of our subgiant-branch sample from the youngest and oldest isochrones (light blue and orange backgrounds, respectively). **c**, **d**, As in **a** (**c**) and **b** (**d**), but for subgiant-branch stars in the cluster core, that is, for stars located at radii of ≤ 20 arcsec.

spread of these stars in the colour-magnitude diagram to derive a maximum likely age spread, Δt , we conclude that $\Delta t \leq 160$ Myr for the full sample and that $\Delta t \leq 80$ Myr for the core sample (Methods).

If the cluster's stellar population were characterized by an age dispersion, this would naturally produce a broadened subgiant branch. Using Fig. 3, we quantitatively assess the probability of the presence of a genuine internal age dispersion. We calculated the number density distributions of both 'typical' stars with extended turn-offs (Fig. 3, blue points) and the cluster's subgiant-branch stars (Fig. 3, red points), adopting differently aged isochrones. The resulting distributions are indeed significantly different, as shown in Fig. 4. Whereas the stars with extended turn-offs exhibit a spread from $\log[t(\text{yr})] = 9.24$ to $\log[t(\text{yr})] = 9.34$, the subgiant-branch stars are almost all associated with the youngest isochrone. Once again, this indicates the lack of a genuine age spread within the cluster.

It is imperative to probe beyond the extended turn-off to fully understand the evolution of massive clusters at ages in excess of 1 Gyr. Subgiant-branch stars will not yet have experienced significant mass loss, which would further complicate our interpretation of, for example, the morphology of the upper end of the red-giant branch and of the red clump, that is, the feature in the Hertzsprung-Russell diagram corresponding to the 'horizontal branch', but for metal-rich stars. Investigation of the subgiant-branch morphology thus offers direct insight into the extent to which intermediate-age clusters resemble true simple stellar populations, unimpeded by effects due to unresolved binary systems^{22,23} or the possible presence of a population of rapidly rotating stars^{8,24–26}, both of which complicate our interpretation of the nature of the observed extended turn-offs. Unresolved binary systems will broaden the turn-off towards lower magnitudes, but they will not cause a reddening of this region²³. Our discovery of a very narrow subgiant branch in NGC 1651 implies that the impact of binary systems is negligible.

The possible presence of a population of rapidly rotating stars may also complicate our interpretation of the observed, extended turn-off regions in intermediate-age clusters^{8,24}. Moreover, because of the conservation of angular momentum, any rapidly rotating stars on the main sequence are (naively) expected to slow quickly when they expand and evolve onto the subgiant branch. However, in practice the contribution

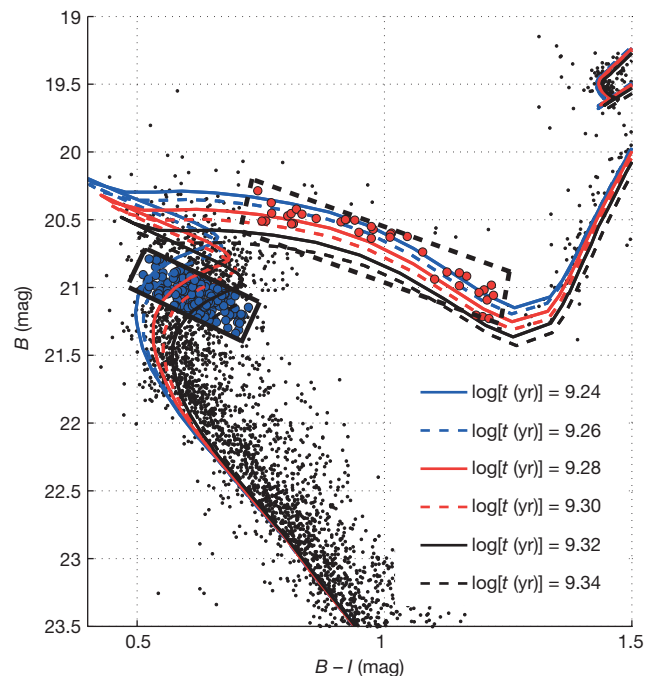


Figure 3 | Comparison of the numbers of stars in NGC 1651 at selected evolutionary stages. Blue points, 'typical' turn-off stars used as basis for the comparison; red points, comparison sample of subgiant-branch stars. Isochrones for different ages are also shown (see key).

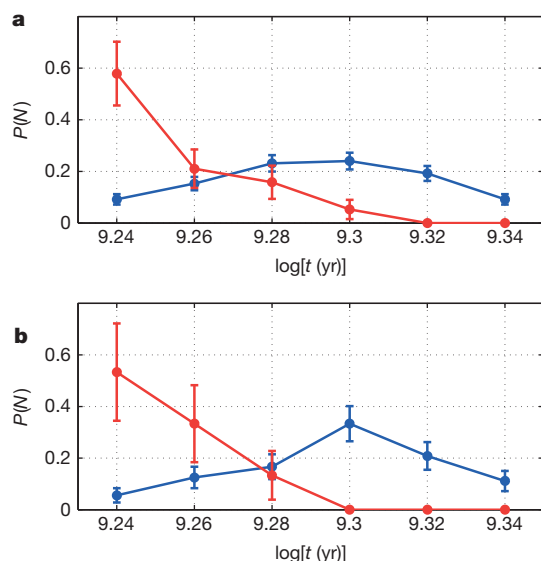


Figure 4 | Expected age distributions resulting from the cluster's turn-off and subgiant-branch stars. **a**, Number density distribution, $P(N)$ (including 1σ standard deviations), as a function of age. Blue, stars in the extended turn-off region; red, subgiant-branch stars. **b**, As in **a**, but for the cluster core region.

to the subgiant-branch morphology from a population of rapidly rotating stars is complex, given that fast stellar rotation leads to longer main-sequence lifetimes²⁷. The presence of such stars may, in fact, also cause a subgiant-branch split²⁸, driven by the resulting extended characteristic stellar mass range and its corresponding range in evolutionary timescale. However, the importance of such a split strongly depends on the prevailing mixing efficiency²⁶. For sufficiently small mixing efficiencies, the turn-off region will be broadened while the subgiant branch will remain relatively narrow (Methods).

Nevertheless, the observed narrow subgiant-branch width provides strong evidence that NGC 1651 cannot have undergone star formation for any significant, sustained length of time. This thus implies that an extended turn-off in the colour–magnitude diagram of an intermediate-age massive cluster does not necessarily imply the presence of a significant, $\gtrsim 100$ Myr age dispersion. NGC 1651 is so far unique, because its subgiant branch is the narrowest yet discovered and discussed for any cluster characterized by an extended turn-off, thus supporting the argument that it is a genuine simple stellar population (for chemical composition-related arguments, see Methods). In retrospect, other intermediate-age clusters have been found that exhibit extended turn-offs but which also exhibit very narrow subgiant branches, including NGC 1783^{2,3,5}, NGC 1806^{2,3,5}, NGC 1846^{2,3,5}, NGC 2155¹¹ and SL 674¹¹. The results highlighted here have left us with an as-yet-unresolved puzzle regarding the evolution of young and intermediate-age massive star clusters. This is troublesome, because star clusters are among the brightest stellar population components in any galaxy; they are visible to much greater distances than are individual stars, even the brightest. Understanding star cluster composition in detail is therefore imperative to understanding the evolution of galaxies as a whole.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 August; accepted 14 October 2014.

- Mackey, A. D. & Broby Nielsen, P. A double main-sequence turn-off in the rich star cluster NGC 1846 in the Large Magellanic Cloud. *Mon. Not. R. Astron. Soc.* **379**, 151–158 (2007).
- Mackey, A. D., Broby Nielsen, P., Ferguson, A. M. N. & Richardson, J. C. Multiple stellar populations in three rich Large Magellanic Cloud star clusters. *Astrophys. J.* **681**, L17–L20 (2008).

- Milone, A. P., Bedin, L. R., Piotto, G. & Anderson, J. Multiple stellar populations in Magellanic Cloud clusters. I. An ordinary feature for intermediate age globulars in the LMC? *Astron. Astrophys.* **497**, 755–771 (2009).
- Rubele, S., Kerber, L. & Girardi, L. The star-formation history of the Small Magellanic Cloud star cluster NGC 419. *Mon. Not. R. Astron. Soc.* **403**, 1156–1164 (2010).
- Goudfrooij, P., Puzia, T. H., Kozhurina-Platais, V. & Chandar, R. Population parameters of intermediate-age star clusters in the Large Magellanic Cloud. II. New insights from extended main-sequence turn-offs in seven star clusters. *Astrophys. J.* **737**, 3 (2011).
- Keller, S. C., Mackey, A. D. & Da Costa, G. S. Extended star formation in the intermediate-age Large Magellanic Cloud star cluster NGC 2209. *Astrophys. J.* **761**, L5 (2012).
- Rubele, S. *et al.* The star formation history of the Large Magellanic Cloud star clusters NGC 1846 and NGC 1783. *Mon. Not. R. Astron. Soc.* **430**, 2774–2788 (2013).
- Li, C., de Grijs, R. & Deng, L. Not-so-simple stellar populations in the intermediate-age Large Magellanic Cloud star clusters NGC 1831 and NGC 1868. *Astrophys. J.* **784**, 157 (2014).
- Bastian, N. & Goodwin, S. P. Evidence for the strong effect of gas removal on the internal dynamics of young stellar clusters. *Mon. Not. R. Astron. Soc.* **369**, L9–L13 (2006).
- Longmore, S. N. *et al.* in *Protostars and Planets VI* (eds Beuther, H., Klessen, R., Dullermond, C. & Henning, Th.) (Univ. Arizona Press, in the press); preprint at <http://arxiv.org/abs/1401.4175> (2014).
- Piatti, A. E., Keller, S. C., Mackey, A. D. & Da Costa, G. S. Gemini/GMOS photometry of intermediate-age star clusters in the Large Magellanic Cloud. *Mon. Not. R. Astron. Soc.* **444**, 1425–1441 (2014).
- Piotto, G. Metallicities on the double main sequence of ω Centauri imply large helium enhancement. *Astrophys. J.* **621**, 777–784 (2005).
- Piotto, G. *et al.* A triple main sequence in the globular cluster NGC 2808. *Astrophys. J.* **661**, L53–L56 (2007).
- Sollima, A. *et al.* Deep FORS1 observations of the double main sequence of ω Centauri. *Astrophys. J.* **654**, 915–922 (2007).
- Milone, A. P. *et al.* The ACS Survey of Galactic Globular Clusters. III. The double subgiant branch of NGC 1851. *Astrophys. J.* **673**, 241–250 (2008).
- Lee, J.-W., Kang, Y.-W., Lee, J. & Lee, Y.-W. Enrichment by supernovae in globular clusters with multiple populations. *Nature* **462**, 480–482 (2009).
- Milone, A. P. *et al.* A double main sequence in the globular cluster NGC 6397. *Astrophys. J.* **745**, 27 (2012).
- Piotto, G. *et al.* Multi-wavelength Hubble Space Telescope photometry of stellar populations in NGC 288. *Astrophys. J.* **775**, 15 (2013).
- Marigo, P. *et al.* Evolution of asymptotic giant branch stars. II. Optical to far-infrared isochrones with improved TP-AGB models. *Astron. Astrophys.* **482**, 883–905 (2008).
- Dirsch, B., Richtler, T., Gieren, W. P. & Hilker, M. Age and metallicity for six LMC clusters and their surrounding field population. *Astron. Astrophys.* **360**, 133–160 (2000).
- Grocholski, A. J., Sarajedini, A., Olsen, K. A. G., Tiede, G. P. & Mancone, C. L. Distances to populous clusters in the Large Magellanic Cloud via the K-band luminosity of the red clump. *Astrophys. J.* **134**, 680–693 (2007).
- Hu, Y., Deng, L., de Grijs, R., Liu, Q. & Goodwin, S. P. The binary fraction of the young cluster NGC 1818 in the Large Magellanic Cloud. *Astrophys. J.* **724**, 649–656 (2010).
- Li, C., de Grijs, R. & Deng, L. The binary fractions in the massive young Large Magellanic Cloud star clusters NGC 1805 and NGC 1818. *Mon. Not. R. Astron. Soc.* **436**, 1497–1512 (2013).
- Bastian, N. & de Mink, S. E. The effect of stellar rotation on colour–magnitude diagrams: on the apparent presence of multiple populations in intermediate age stellar clusters. *Mon. Not. R. Astron. Soc.* **398**, L11–L15 (2009).
- Li, Z., Mao, C., Chen, L. & Zhang, Q. Combined effects of binaries and stellar rotation on the color–magnitude diagrams of intermediate-age star clusters. *Astrophys. J.* **761**, L22 (2012).
- Yang, W., Bi, S., Meng, X. & Liu, Z. The effects of rotation on the main-sequence turn-off of intermediate-age massive star clusters. *Astrophys. J.* **776**, 112 (2013).
- Girardi, L., Eggenberger, P. & Miglio, A. Can rotation explain the multiple main-sequence turn-offs of Magellanic Cloud star clusters? *Mon. Not. R. Astron. Soc.* **412**, L103–L107 (2011).
- Georgy, C. *et al.* Populations of rotating stars. III. SYCLIST, the new Geneva population synthesis code. *Astron. Astrophys.* **566**, A21 (2014).

Acknowledgements We thank S. de Mink, Y. Huang and X. Chen for discussions and assistance. Partial financial support for this work was provided by the National Natural Science Foundation of China through grants 11073001, 11373010 and 11473037.

Author Contributions C.L., R.d.G. and L.D. jointly designed and coordinated this study. C.L. performed the data reduction. C.L. and R.d.G. collaborated on the detailed analysis. L.D. provided ideas that improved the study's robustness. All authors read, commented on and jointly approved submission of this article.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.L. (joshuali@pku.edu.cn).

METHODS

Observations and data reduction. The data sets of NGC 1651 were obtained as part of Hubble Space Telescope programme GO-12257 (principal investigator: L. Girardi), using the Wide Field Camera 3 (WFC3). Both clusters were observed through the F475W and F814W filters (with central wavelengths of 475 nm and 814 nm, respectively), which roughly correspond to the Johnson–Cousins B and I bands, respectively. Two images with long exposure times of 1,440 s and 1,430 s in the B and I bands, respectively, in addition to two images with short exposure times of 720 s and 700 s, respectively, were obtained. We used the IRAF/DAOPHOT software package to perform point-spread-function photometry²⁹.

The photometric catalogues pertaining to the long- and short-exposure-time images were combined. We carefully cross-referenced both catalogues to avoid duplication of objects in the combined output catalogue. For stars in common between both catalogues, we adopted the generally more accurate photometry from the long-exposure catalogue for inclusion in the output master catalogue, except in the magnitude range where the long-exposure image could be affected by saturation²³ (for example for stars on the upper red-giant branch or blue stragglers).

Determination of the cluster region. We divided the stellar spatial distribution into 20 bins along both the right ascension (α_{J2000}) and declination (δ_{J2000}) axes. Using a Gaussian function to fit the stellar number density distribution in each direction, we determined the closest coincidence of both Gaussian peaks as the cluster centre: $\alpha_{J2000} = 04\text{ h }37\text{ min }32.16\text{ s}$ (69.3843°), $\delta_{J2000} = -70^\circ 35' 08.88''$ (-70.5858°). The centre position compares very well with previous determinations. For instance, NASA's Extragalactic Database (<http://ned.ipac.caltech.edu>) lists $\alpha_{J2000} = 04\text{ h }37\text{ min }32.3\text{ s}$, $\delta_{J2000} = -70^\circ 35' 9''$, and the Strasbourg Astronomical Data Center's SIMBAD (<http://simbad.u-strasbg.fr/>) gives $\alpha_{J2000} = 04\text{ h }37\text{ min }31.1\text{ s}$, $\delta_{J2000} = -70^\circ 35' 2''$, compared with the NGC/IC Project's (<http://www.ngcicproject.org/realskyview/N1600-N1699.txt>) $\alpha_{J2000} = 4.625750\text{ h} \equiv 69.38625^\circ$, $\delta_{J2000} = -70.585560^\circ$.

The complete data set for this cluster is composed of a combination of two WFC3 images. We used a Monte Carlo-based method to estimate the areas of rings of different radii (all radii were measured from the centre of the cluster). Specifically, we calculated the total area of the region covered and subsequently generated millions of points that were homogeneously distributed across the full region. We then calculated the number of points located in each ring as a fraction of the total number of points. We used this fraction, multiplied by the total area, to represent the specific area of each ring. The number of stars in each ring is $N(R)/A(R)$, where $N(R)$ is the number of observed stars located in a ring with radius R and $A(R)$ is the corresponding area of the ring.

We next calculated the total brightness of stars in each ring, $f(R) = \sum_N 10^{(B_N - (m - M)_0)/(-2.5)}$, where N is the number of stars located in the ring of interest; B is the B-band magnitude and its subscript N refers to the running number of the summation; and $(m - M)_0 = 18.46\text{ mag}$ is the adopted distance modulus. The brightness density is then $\rho_f(R) = f(R)/A(R)$, which corresponds to a surface brightness of $\mu(R) = -2.5\log[\rho_f(R)] + 18.46$. Because NGC 1651 is an intermediate-age star cluster, we represent its brightness profile by^{30,31}

$$f(r) = f_0 \left(1 + \frac{r^2}{a^2}\right)^{-\gamma/2}$$

where f_0 is the central surface brightness. The measures of the core radius, a , and the power-law index, γ , are linked to the King core radius, r_c , through

$$r_c = a(2^{2/\gamma} - 1)^{1/2}$$

The cluster's radial profile, including the 1σ photometric uncertainties due to Poisson noise, as well as the best-fitting theoretical profile, are shown in Extended Data Fig. 1.

Field-star decontamination. The Hubble Space Telescope/WFC3 images cover a very large region, allowing us to investigate the entire cluster as well as a neighbouring field region. On the basis of the radial density profile in Extended Data Fig. 2, we determined that for $R \geq 85\text{ arcsec}$ the cluster brightness becomes indistinguishable from the background noise. We hence selected the region characterized by $R \geq 85\text{ arcsec}$ as our comparison field region for the purposes of field-star decontamination. Taking into account the standard deviation of the field-star magnitudes, we concluded that the most representative cluster region has a radius of $R = 75\text{ arcsec}$. We statistically field-star decontaminated this cluster region. Using a Monte Carlo approach, we estimated that the comparison field region covers 46.7% of the cluster region.

The full stellar catalogue resulting from our analysis of the field region contains 759 stars. Given that the cluster region covers 2.14 times that of the comparison field, from a statistical perspective we expect 1,607 field stars to be located within the cluster region. We divided the NGC 1651 cluster and field colour–magnitude diagrams into 50 bins in magnitude and 25 bins in colour; for relatively sparsely

populated regions, we enlarged the bin size appropriately (see below). We then calculated the number of field stars in each colour–magnitude bin, and subsequently removed 2.14 times the (integer) number of stars from the corresponding bins of the NGC 1651 colour–magnitude diagram.

Because the comparison field region was selected from the same image as the cluster region, its exposure time is identical. Hence, exposure-time differences will not affect the reliability of our field-star decontamination, although statistical differences between the cluster and field regions cause a slight dependence on the adopted grid size. We carefully checked how the number of bins adopted would affect the decontamination results and enlarged the bin sizes for sparsely populated regions (for example on the red side of the main sequence). We concluded that our field-star decontamination is robust with respect to reasonable differences in adopted bin size. This thus eventually resulted in a statistically robustly field-star-decontaminated colour–magnitude diagram of NGC 1651. We show the results of the main steps used in our field-star decontamination procedure in Extended Data Fig. 2. Extended Data Fig. 2a shows the original colour–magnitude diagram of NGC 1651 (for $R \leq 75\text{ arcsec}$), Extended Data Fig. 2b represents the synthesized field-star equivalent and Extended Data Fig. 2c is the decontaminated colour–magnitude diagram on which we based our analysis.

Using the subgiant branch to constrain the cluster's maximum age dispersion.

Many authors have invoked age dispersions to explain the observed extended turn-off regions, and although numerous, apparently somewhat different scenarios have been proposed, most can be traced back to the basic idea of an age dispersion. For instance, mergers of star clusters with an age difference of $\sim 200\text{ Myr}$ (ref. 1), as well as interactions of star clusters and star-forming giant molecular clouds³², have been suggested as the possible origin of extended turn-off regions.

We calculated the magnitude deviation (ΔB) with respect to the youngest ($\log[t(\text{yr})] = 9.24$) isochrone for each subgiant-branch star (Fig. 2). Because our full sample contains 38 subgiant-branch stars, we adopted five bins in ΔB . A gradually increasing trend in ΔB is found, starting from $\Delta B \approx -0.09\text{ mag}$, with a peak at $\Delta B \approx 0.00\text{ mag}$, followed by a decrease to $\Delta B \approx 0.14\text{ mag}$ and a slight upturn to $\Delta B \approx 0.20\text{ mag}$: see Extended Data Fig. 3, which includes the 1σ standard deviations. We next generated an additional set of isochrones characterized by different ages and applied the same procedure. The typical ΔB values are included at the top of Extended Data Fig. 3 (black dashed lines), for an age resolution of $\Delta\log[t(\text{yr})] = 0.2$.

Assuming appropriate photometric uncertainties for each of these isochrones, we calculated the number of subgiant-branch stars that would be covered if we adopted a given age dispersion. We first proceeded to test the simple-stellar-population approximation, that is, assuming no age dispersion. In this case, all stars should be located on the $\log[t(\text{yr})] = 9.24$ isochrone, with a spread determined by the typical (3σ) photometric uncertainties of 0.12 mag . We found that 30 of our 38 subgiant-branch stars are associated with this isochrone (Fig. 3, light blue background).

If we assume that all subgiant-branch stars belong to a simple stellar population characterized by a typical age of $\log[t(\text{yr})] = 9.26$, and adopting the same photometric uncertainties, we can reproduce 35 of the 38 stars (92%). This thus strongly implies that the NGC 1651 stellar population is most probably a genuine simple stellar population. Extended Data Table 1 includes the results of our analysis to derive the maximum intrinsic age dispersion needed to explain the observed subgiant-branch loci in the cluster's colour–magnitude diagram.

An age dispersion of $\sim 80\text{ Myr}$ can reproduce $>90\%$ of the subgiant-branch stars in our full sample. Similarly, if we assume that the cluster's subgiant-branch stars are members of a simple stellar population, a typical age of $\log[t(\text{yr})] = 9.26$ can also reproduce $>90\%$ of all subgiant-branch stars. The result holds for the subgiant-branch sample in the cluster core: an age dispersion of $\sim 80\text{ Myr}$ can reproduce all the core subgiant-branch stars, and a simple-stellar-population model with a typical age of $\log[t(\text{yr})] = 9.26$ still reproduces $>90\%$ of the core subgiant-branch stars. This hence unequivocally excludes the presence of an age dispersion extending to at least $\log[t(\text{yr})] = 9.34$.

A population of rapidly rotating stars? The observed extended turn-off regions in intermediate-age clusters might also be explained as evidence of the presence of a population of rapidly rotating stars^{8,24,26}. The centrifugal force resulting from rapid stellar rotation leads to a reduction in effective gravity, which decreases both the stellar surface temperature and its luminosity²⁶. The reduced gravity also leads to a decreasing stellar central hydrogen-burning efficiency, rendering stars slightly fainter. This effect mainly affects F-type stars; stars with masses below 1.2 solar masses do not rotate rapidly, because of magnetic braking³³.

Although some authors have claimed that rapid stellar rotation could lead to a broadening of the turn-off^{8,25}, this scenario holds only if rapid rotation does not have any effect on the stellar lifetime on the main sequence. However, rapid rotation will also cause a transfer of mass from radial shells to the central core, thus providing additional material for nuclear fusion in the core. This could increase the lifetimes

of rotating stars relative to those of their non-rotating counterparts. Calculations of the effect of this expected prolongation of stellar lifetimes have led some authors²⁷ to conclude that the resulting colour–magnitude diagram will still retain a narrow turn-off. These authors maintained that the presence of an age dispersion was still the most natural model that reproduces the extended turn-off. However, derivation of colour–magnitude diagrams resulting from the adoption of different rotation velocities²⁶, while also considering the increased main-sequence lifetimes, led to the conclusion that such a scenario can still reproduce the observed extended turn-offs. However, the extent of the turn-off's broadening depends on the typical cluster age. Nevertheless, if one adopts a modest mixing efficiency for rotating stars, extended turn-offs can still be observed²⁶. In any case, because different stellar rotation rates have been observed for solar-neighbourhood field stars³⁴, it is natural to expect that stars in star clusters may have similar distributions of rotation velocities.

Overall, the extent to which rapid rotation will affect subgiant-branch stars is as yet unclear. Very few authors consider these effects, with the exception of a single study²⁸ that aims to generate a grid of stellar models including a range of rotation rates. Although these authors have thus far only satisfactorily completed their calculations for extremely massive stars, using different evolutionary tracks and a range of rotation velocities, this allows us to estimate the extent to which rapid rotation may affect stars on the subgiant branch. On the basis of their interactive tools (<http://obswww.unige.ch/Recherche/evol/-Database->), we generated two evolutionary tracks for their lowest-mass stars, each of 1.7 solar masses, one without rotation and the other characterized by extremely rapid rotation ($\omega = 0.95$, that is, rotation at 95% of the critical break-up rate): see Extended Data Fig. 4. We see that, following the turn-off stage, the rapidly rotating track converges to the non-rotating track. Indeed, because of the conservation of angular momentum, the fast rotators are expected to slow quickly when they expand and evolve onto the subgiant branch. This result hence confirms that the effects of rapid stellar rotation become negligible, such that the observed narrow subgiant branch in NGC 1651 can be reconciled only with the colour–magnitude diagram of a genuine simple stellar population.

However, taking into account the effects of rapid rotation is highly complex. Because stars that originally rotate rapidly tend to live longer than their non-rotating counterparts, the presence of a population of rapidly rotating stars may, in fact, still give rise to a broadened or split subgiant branch. Whether or not this scenario holds depends on the atmospheric mixing efficiency, the effects of which are as yet unclear. Nevertheless, we point out that if the mixing efficiency is reduced to 'normal' levels of 0.03, the extended turn-off caused by the most rapid stellar rotation will be equivalent to a simple-stellar-population age spread of approximately 450 Myr for clusters aged 1.7 Gyr (ref. 26). This fits our observations exactly.

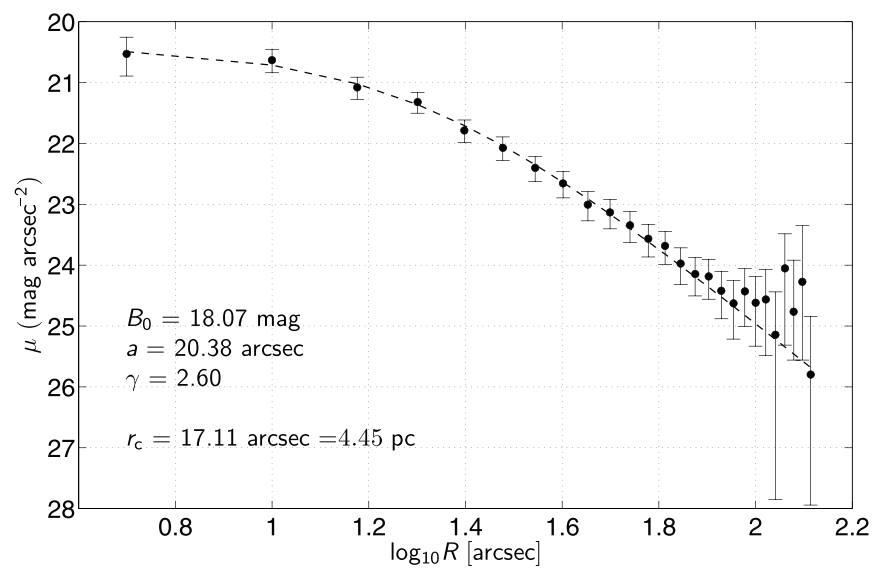
Additional evidence in support of NGC 1651 as a simple stellar population.

Except for possibly the cluster's sodium abundance, [Na/Fe], the observed dispersions in the abundances of all other elements investigated so far are consistent with the measurements' root-mean-squared values³⁵. [Na/Fe] ranges from approximately -0.41 dex to -0.03 dex, but this result is based on analysis of only five bright asymptotic-giant-branch stars, which may be strongly affected by their associated stellar winds. In fact, it has been shown convincingly^{35,36} that a number of clusters with extended turn-offs do not exhibit chemical-abundance spreads.

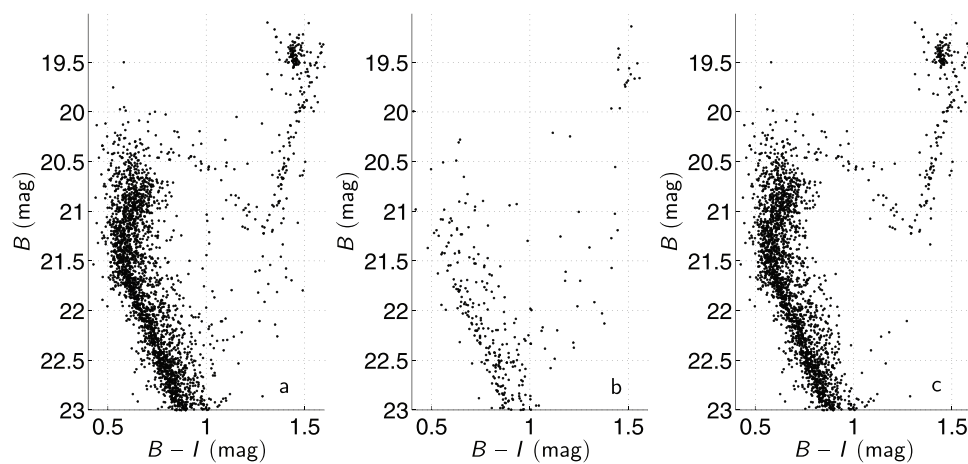
On the basis of a detailed analysis of the spectra of 1,200 red giants in 19 clusters³⁷, it has become apparent that first-, intermediate- and extreme second-generation stars tend to be found in three typical zones in the [Na/Fe]–[O/Fe] diagram. In this context, first-generation stars may be characterized by relatively poor sodium abundances, exhibiting dispersions of up to 0.4 dex (ref. 37). Therefore, the absence of any significant abundance dispersions in most elements³⁵ in the cluster, combined with the observed spread in [Na/Fe], is indeed consistent with NGC 1651 representing a genuine simple stellar population.

Recent insights³⁸ convincingly showed that star clusters with ages of up to 300 Myr in both Magellanic Clouds do not have any sizeable gas reservoirs left to form second-generation stars. One must thus turn to alternative models to explain the observations of clusters like NGC 1651.

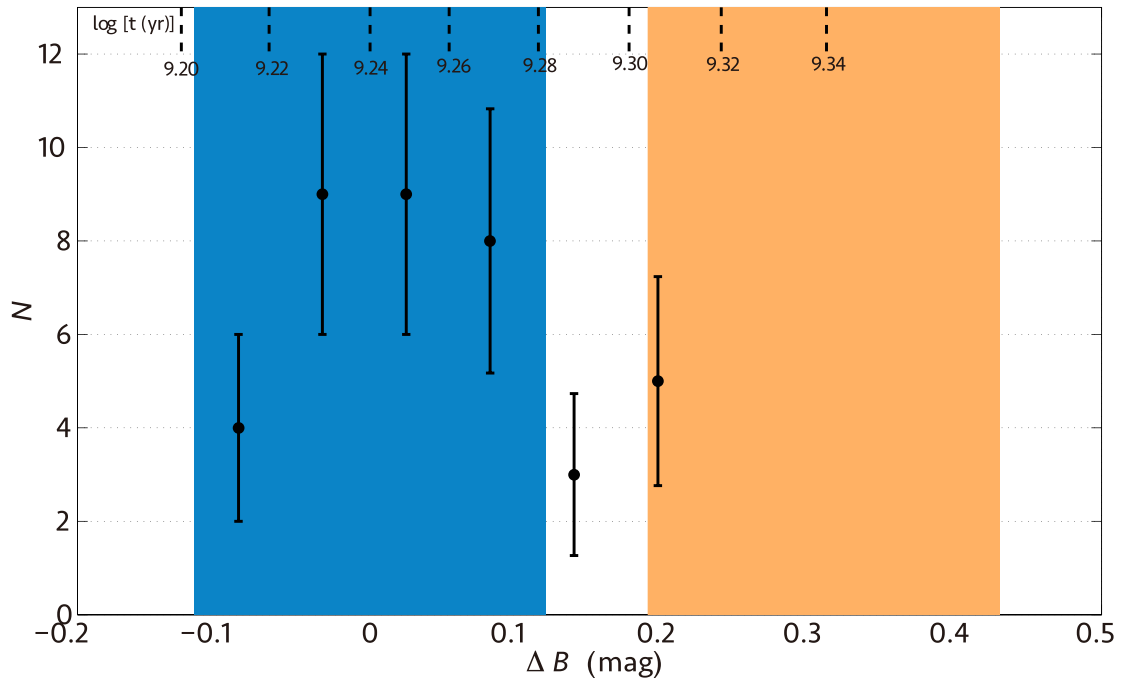
29. Davis, L. E. A Reference Guide to the IRAF/DAOPHOT Package. <http://iraf.noao.edu/iraf/web/docs/recommend.html> (1994).
30. Elson, R. A. W., Fall, S. M. & Freeman, K. C. The stellar content of rich young clusters in the Large Magellanic Cloud. *Astrophys. J.* **336**, 734–751 (1989).
31. Mackey, A. D. & Gilmore, G. F. Surface brightness profiles and structural parameters for 53 rich stellar clusters in the Large Magellanic Cloud. *Mon. Not. R. Astron. Soc.* **338**, 85–119 (2003).
32. Bekki, K. & Mackey, A. D. On the origin of double main-sequence turn-offs in star clusters of the Magellanic Clouds. *Mon. Not. R. Astron. Soc.* **394**, 124–132 (2009).
33. Mestel, L. & Spruit, H. C. On magnetic braking of late-type stars. *Mon. Not. R. Astron. Soc.* **226**, 57–66 (1987).
34. Royer, F., Zorec, J. & Gómez, A. E. Rotational velocities of A-type stars. III. Velocity distributions. *Astron. Astrophys.* **463**, 671–682 (2007).
35. Mucciarelli, A., Carretta, E., Origlia, L. & Ferraro, F. R. The chemical composition of red giant stars in four intermediate-age clusters of the Large Magellanic Cloud. *Astron. J.* **136**, 375–388 (2008).
36. Mucciarelli, A., Dalessandro, E., Ferraro, F. R., Origlia, L. & Lanzoni, B. No evidence of chemical anomalies in the bimodal turnoff cluster NGC 1806 in the Large Magellanic Cloud. *Astrophys. J.* **793**, L6 (2014).
37. Carretta, E. *et al.* Properties of stellar generations in globular clusters and relations with global parameters. *Astron. Astrophys.* **516**, A55 (2010).
38. Bastian, N. & Strader, J. Constraining globular cluster formation through studies of young massive clusters. III. A lack of gas and dust in massive stellar clusters in the LMC and SMC. *Mon. Not. R. Astron. Soc.* **443**, 3594–3600 (2014).



Extended Data Figure 1 | Radial brightness density profile of NGC 1651. The 1σ uncertainties shown are due to Poisson noise.

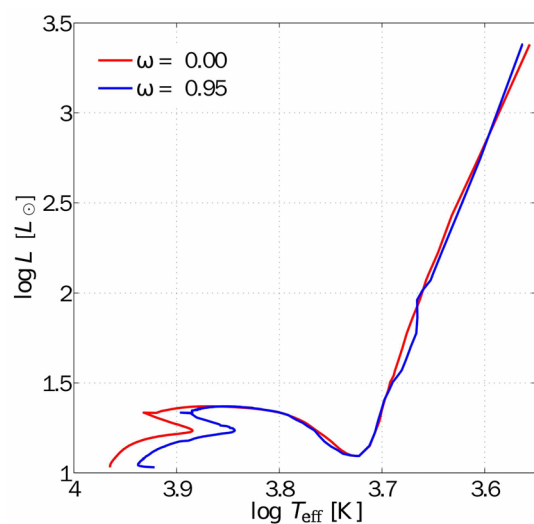


Extended Data Figure 2 | Background decontamination. **a**, Original colour-magnitude diagram of NGC 1651. **b**, Field-star colour-magnitude diagram. **c**, Field-star-decontaminated NGC 1651 colour-magnitude diagram.



Extended Data Figure 3 | Constraints on the maximum likely age dispersion. Number distribution, N (including 1σ standard deviations), of the deviations in magnitude, ΔB , of our subgiant-branch sample, as in Fig. 2. The

black dashed lines at the top indicate typical ΔB values for isochrones of different ages, as indicated.



Extended Data Figure 4 | Evolutionary tracks for extremes in stellar rotation rates. Red, non-rotating stars; blue, stellar rotation at 95% of the critical break-up rate ($\omega = 0.95$). Both tracks apply to 1.7 solar-mass stars. L_{\odot} , solar luminosity; T_{eff} , effective temperature.

Extended Data Table 1 | Age dispersions required to match the observed spread of subgiant-branch stars in NGC 1651

$\Delta\log[t(\text{yr})]$	N_{SGB}	Fraction (%)	$\Delta t(\text{Myr})$
9.24–9.28	38/38	100.0	167
9.26–9.28	37/38	97.4	86
9.24–9.26	36/38	94.7	82
9.26	35/38	92.1	SSP
9.24	30/38	78.9	SSP
9.28	27/38	71.1	SSP
9.24–9.26	15/15	100.0	82
9.26	14/15	93.3	SSP
9.24	13/15	86.7	SSP
9.28	10/15	66.7	SSP

$\Delta\log(t)$, age dispersion. Top, full sample; bottom, subgiant-branch stars in the cluster core. SGB, subgiant branch; SSP, simple stellar population.

Deterministic switching of ferromagnetism at room temperature using an electric field

J. T. Heron¹, J. L. Bosse², Q. He³, Y. Gao^{4,5}, M. Trassin⁶, L. Ye², J. D. Clarkson⁷, C. Wang⁸, Jian Liu⁴, S. Salahuddin⁹, D. C. Ralph^{8,10}, D. G. Schlom^{1,10}, J. Íñiguez¹¹, B. D. Huey^{2,12} & R. Ramesh^{4,7,13}

The technological appeal of multiferroics is the ability to control magnetism with electric field^{1–3}. For devices to be useful, such control must be achieved at room temperature. The only single-phase multiferroic material exhibiting unambiguous magnetoelectric coupling at room temperature is BiFeO₃ (refs 4 and 5). Its weak ferromagnetism arises from the canting of the antiferromagnetically aligned spins by the Dzyaloshinskii–Moriya (DM) interaction^{6–9}. Prior theory considered the symmetry of the thermodynamic ground state and concluded that direct 180-degree switching of the DM vector by the ferroelectric polarization was forbidden^{10,11}. Instead, we examined the kinetics of the switching process, something not considered previously in theoretical work^{10–12}. Here we show a deterministic reversal of the DM vector and canted moment using an electric field at room temperature. First-principles calculations reveal that the switching kinetics favours a two-step switching process. In each step the DM vector and polarization are coupled and 180-degree deterministic switching of magnetization hence becomes possible, in agreement with experimental observation. We exploit this switching to demonstrate energy-efficient control of a spin-valve device at room temperature. The energy per unit area required is approximately an order of magnitude less than that needed for spin-transfer torque switching^{13,14}. Given that the DM interaction is fundamental to single-phase multiferroics and magnetoelectrics^{3,9}, our results suggest ways to engineer magnetoelectric switching and tailor technologically pertinent functionality for nanometre-scale, low-energy-consumption, non-volatile magnetoelectronics.

The ability of magnetoelectric multiferroics^{2,3} to couple magnetic and ferroelectric orders suggests that they have the potential to add functionality to devices while also reducing energy consumption¹⁵. BiFeO₃ is the only thermodynamically stable room-temperature magnetoelectric multiferroic material^{16–19}. Understanding the mechanisms operating in magnetoelectric multiferroics and engineering their properties is imperative if we are to address this lack of room-temperature functionality⁴. In BiFeO₃, as well as many other multiferroics^{3,8,9}, an antisymmetric exchange (DM) interaction manifests from spin–orbit coupling and this interaction gives rise to a weak ferromagnetic moment through the canting of the antiferromagnetically aligned Fe³⁺ spins⁹. From thermodynamic considerations, which inherently assume that polarization switching occurs in a single step, it was concluded that it would be impossible to achieve deterministic 180° electric-field-induced switching of the weak ferromagnetism, without some other change in the system, such as a change in the sense of the oxygen octahedral rotation^{10,11}. Recently, theorists have started to consider the kinetics of the switching path in multiferroic switching; however, the focus has been on multiferroics other than BiFeO₃ (ref. 20). First-principles calculations of

BiFeO₃-based systems have also predicted novel magnetoelectric coupling mechanisms that lead to electric-field control of magnetism^{21,22}.

Here we present a combined experimental and theoretical study where the DM vector (defined later in the discussion of Fig. 2) and weak ferromagnetism of a strained BiFeO₃ film switch by 180° through the application of an out-of-plane electric field. Driven by a two-step sequential rotation of the polarization vector upon application of the electric field, the oxygen octahedral rotations (which determine the DM vector and weak magnetization) follow this two-step sequence, leading to the reversal of the DM vector and weak magnetization. Our calculations find a large kinetic barrier to single-step switching, making a two-step switching path the preferred path and key to enabling the reversal of the weak ferromagnetism with an electric field. The applicability of such a switching event is demonstrated with the electric-field control of a spin-valve device at room temperature.

(001)_p BiFeO₃ (100 nm)/(001)_p SrRuO₃ (8 nm) heterostructures, where the subscript ‘p’ denotes pseudocubic perovskite indices, were grown on (110)-oriented DyScO₃ substrates, providing a small anisotropic strain to the BiFeO₃ film (see Methods for synthesis details). SrRuO₃ acts as a conductive back-electrode for the application of an out-of-plane electric field for polarization switching. Time-dependent, dual-frequency piezoresponse force microscopy (PFM) studies were used to investigate discrete steps in the polarization switching as a function of position (Methods).

Vector PFM images of the initial (pre-switched) and final states (Fig. 1a) show the BiFeO₃ film to have nominally two polarization domains, which form a stripe-like structure before and after switching. The electric field reverses the $[\bar{1}\bar{1}1]_p$ (brown) and $[1\bar{1}1]_p$ (orange) oriented polarizations to the $[11\bar{1}]_p$ (dark blue) and $[\bar{1}\bar{1}1]_p$ (light blue) directions (Fig. 1a). A circular region near the centre of the initial image was intentionally pre-switched to determine when switching was complete during imaging. Interestingly, multiple polarization orientation changes per location were observed between the initial and final states (Fig. 1b). Over 87% of the switched area experienced multiple orientation changes, with two orientation changes being the most common (Fig. 1c).

The first switch consists of a mixture of in-plane 71° and out-of-plane 109° switches, comprising 55.4% and 40.7% of the total observed switching events, respectively (Fig. 1d; Methods). More than 75% of the area underwent a net switch of 180° as a result of the entire switching sequence. Similar statistics are observed for the return switching process, where the switched region is reverted back into the initial state with a reversed electric field (Extended Data Fig. 2). These data illustrate that an out-of-plane electric field reverses the local polarization in these BiFeO₃ films and that the switching is not direct. Rather, it consists of a two-step switching process that begins with an in-plane 71° or out-of-plane 109° switch and ends with the other.

¹Department of Materials Science and Engineering, Cornell University, Ithaca, New York 14853, USA. ²Department of Materials Science and Engineering, University of Connecticut, Storrs, Connecticut 06269, USA. ³Department of Physics, Durham University, Durham DH1 3LE, UK. ⁴Department of Physics, University of California, Berkeley, California 94720, USA. ⁵School of Materials Science and Engineering, and State Key Lab of New Ceramics and Fine Processing, Tsinghua University, Beijing 100084, China. ⁶Department of Materials, ETH Zurich, Vladimir-Prelog-Weg 4 10, 8093 Zurich, Switzerland. ⁷Department of Materials Science and Engineering, University of California, Berkeley, California 94720, USA. ⁸Department of Physics, Cornell University, Ithaca, New York 14853, USA. ⁹Department of Electrical Engineering and Computer Science, University of California, Berkeley, California 94720, USA. ¹⁰Kavli Institute at Cornell for Nanoscale Science, Ithaca, New York 14853, USA. ¹¹Institut de Ciència de Materials de Barcelona (ICMAB-CSIC), Campus UAB, 08193 Bellaterra, Spain. ¹²Institute of Materials Science, University of Connecticut, Storrs, Connecticut 06269, USA. ¹³Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.

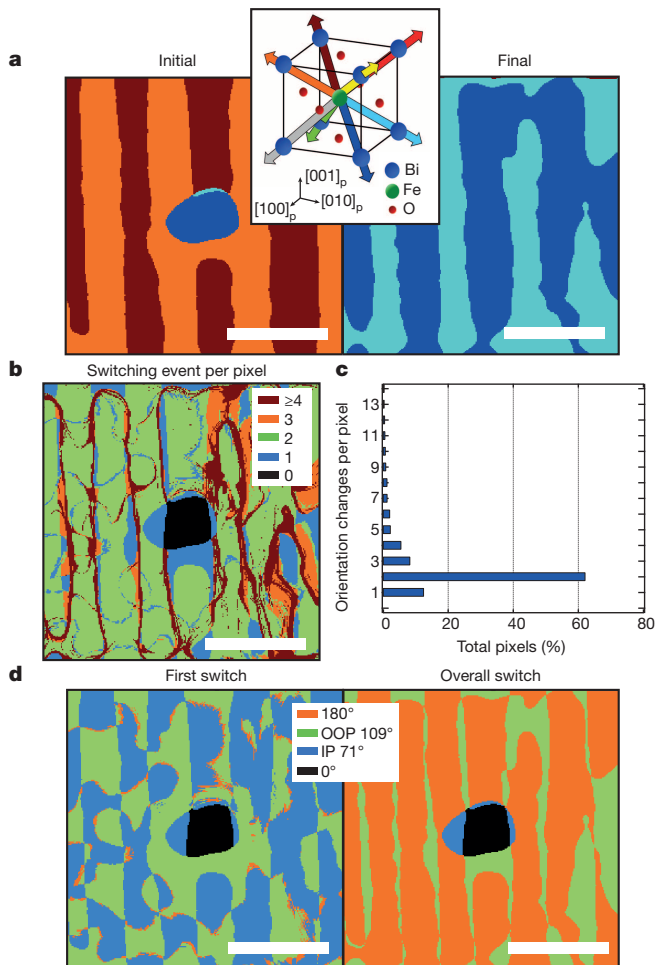


Figure 1 | Two-step polarization switching. **a**, Polarization vector images of initial (pre-switched) and final states of $\text{BiFeO}_3 (001)_p/\text{SrRuO}_3 (001)_p/\text{DyScO}_3 (110)$. A circular region in the initial image was intentionally pre-switched. The inset shows a schematic of the eight possible polarization orientations of BiFeO_3 . **b**, Map of the total number of polarization orientation changes per location through switching. **c**, Number of polarization orientation changes per pixel versus percentage of total pixels (that is, areal percentage). **d**, Maps of the switching directions made by the first and overall reorientations. Scale bars are 500 nm.

Ab initio calculations corroborate this switching scenario. Figure 2a shows the lowest-energy switching path calculated (and optimized with the ‘nudged elastic band’ method; see Methods) with mechanical boundary conditions set by the DyScO_3 substrate; this path is compatible with the experimentally observed two-step polarization switching process (see Methods). The displacement of the bismuth ion in the polarization direction (results for the iron shift are qualitatively identical (Methods)) shows a two-step switching sequence composed of a 71° switch followed by a 109° switch. Interestingly, we find that the axis of rotation of the oxygen octahedra (O_6) follows the bismuth displacement. The trajectory of the direct, single-step switching path is such that the Bi^{3+} shift reverses following a trajectory directly through zero (that is, through zero polarization in contrast to the two-step switching path where the polarization remains non-zero throughout). Ultimately, single-step switching leaves the oxygen octahedra unperturbed and is unlikely to occur because of the large (about 240 meV per formula unit) energy barrier (green curves in Fig. 2a). The prediction of reversal of the O_6 -rotation through the two-step switching has notable implications for the switching of the DM vector and the weak ferromagnetic moment.

In pseudocubic perovskites the DM vector is given by $\mathbf{D} \approx \sum \mathbf{d}_i \times \mathbf{r}_i$, where \mathbf{r}_i is the vector connecting neighbouring Fe^{3+} ions and \mathbf{d}_i is the displacement of the intermediate oxygen atom from the mid-point of \mathbf{r}_i caused by the O_6 -rotation. Since \mathbf{r}_i remains fixed, the reversal of \mathbf{d}_i through two-step polarization switching, switches the sign of \mathbf{D} . Given that the in-plane 71° and out-of-plane 109° switches (Fig. 1d) are magnetoelectric^{23,24}, capable of rotating \mathbf{D} and the associated weak ferromagnetism \mathbf{M}_c by 90° , the reversal of \mathbf{D} through sequential in-plane 71° and out-of-plane 109° switches would also be expected to reverse

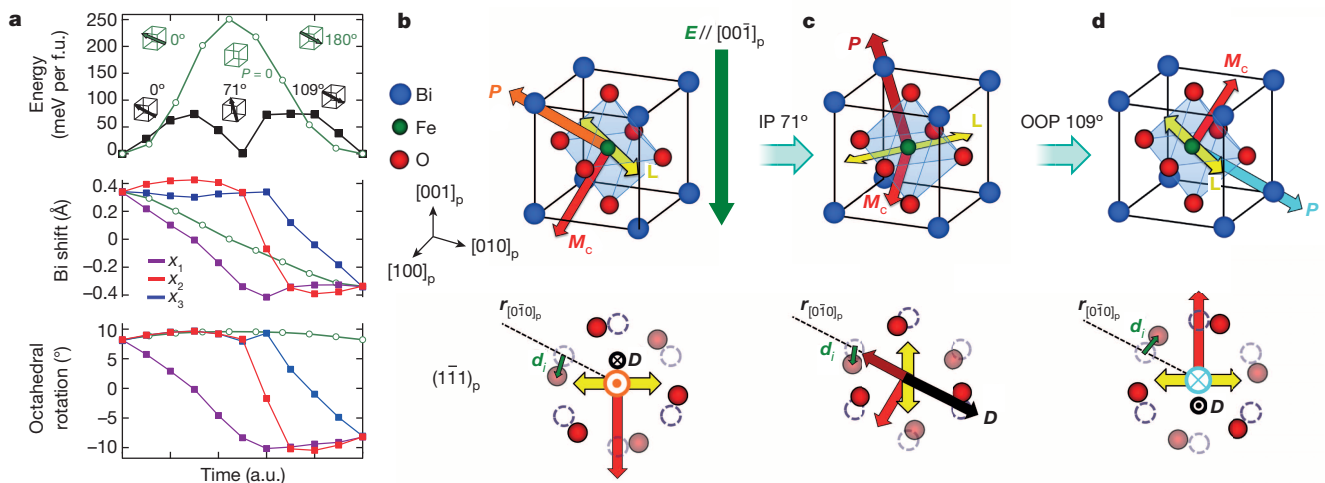


Figure 2 | Magnetoelectric switching path. **a**, Calculated energy (upper panel), off-centre bismuth displacement (middle panel, representing the ferroelectric polarization), and oxygen octahedral (O_6) rotation (lower panel) versus a time coordinate indicating the evolution along the switching trajectory. Results are shown for one-step (direct) and two-step 180° -switching trajectories (schematically illustrated). For the two-step switching path, the black curve shows the energy landscape while purple, blue and red curves indicate orthogonal components (x_1 , x_2 and x_3) of the bismuth displacement and O_6 rotation. Green curves show those for direct 180° ferroelectric switching. f.u., formula unit. a.u., arbitrary unit. **b–d**, Schematics of switching

path viewed in three dimensions (top row) and from the $(1\bar{1}1)_p$ surface (lower). Bismuth shifts are not shown. **b**, A domain with polarization \mathbf{P} initially along the $[111]_p$. \mathbf{L} represents the antiferromagnetic axis. O_6 rotations are shown with respect to an octahedron without rotations (dashed circles). **c**, After an in-plane 71° switch. \mathbf{E} represents the applied electric field. **d**, After an out-of-plane 109° switch. Oxygen displacement \mathbf{d}_i (green arrow) and the line connecting neighbouring Fe^{3+} atoms \mathbf{r}_i are shown for the $[0\bar{1}0]_p$ direction while \mathbf{D} (black arrow) is given for the sum of neighbouring Fe^{3+} atoms. Crossed and dotted circles indicate vectors that point into and out of the page, respectively.

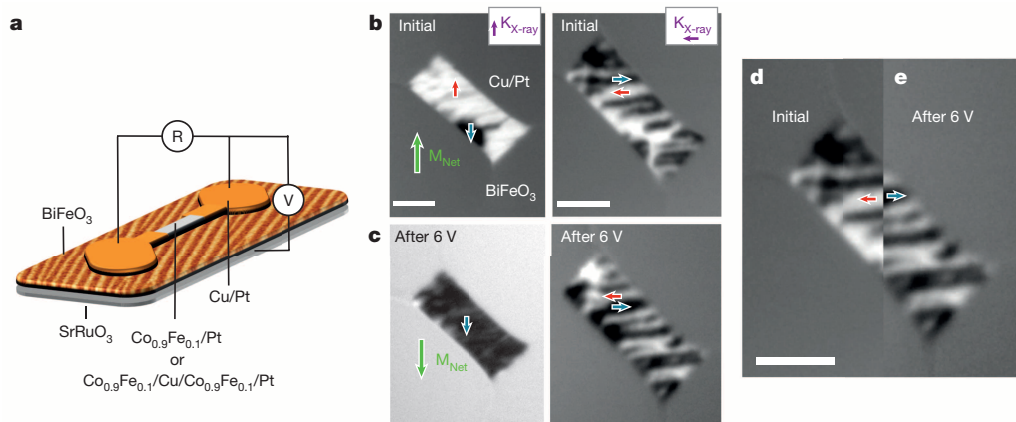


Figure 3 | Magnetization reversal by electric field. **a**, Schematic of a magnetoelectric device consisting of either $\text{Co}_{0.9}\text{Fe}_{0.1}$ or a $\text{Co}_{0.9}\text{Fe}_{0.1}/\text{Cu}/\text{Co}_{0.9}\text{Fe}_{0.1}$ spin valve on BiFeO_3 . Initial **(b)** and final **(c)** directions of the in-plane $\text{Co}_{0.9}\text{Fe}_{0.1}$ moments with components viewed perpendicular (vertical $K_{X\text{-ray}}$, where $K_{X\text{-ray}}$ defines the in-plane component of the incident X-ray beam) and parallel to the stripe domains (horizontal $K_{X\text{-ray}}$). The directions of the magnetizations in each domain are highlighted with blue and

red arrows, which correspond to the local moment direction being perpendicular or parallel to $K_{X\text{-ray}}$. The net $\text{Co}_{0.9}\text{Fe}_{0.1}$ magnetization (green arrows) reverses after the voltage is applied. $K_{X\text{-ray}}$ images of the initial **(d)** and 6 V **(e)** states merged near the centre of the $\text{Co}_{0.9}\text{Fe}_{0.1}$ to reveal the magnetization reversal at each domain. The black defect in **b** (vertical $K_{X\text{-ray}}$) hinders switching in that region. Scale bars are 2 μm .

M_c . The schematics in Fig. 2b–d illustrate how the switching pathway leads to a reversal of D and M_c .

Verification of these theoretical predictions comes from experiments using a heterostructure that consists of a strong ferromagnet, $\text{Co}_{0.9}\text{Fe}_{0.1}$, deposited onto BiFeO_3 . Prior work has shown a one-to-one correlation between the $\text{Co}_{0.9}\text{Fe}_{0.1}$ domain structure and the ferroelectric domain structure of the underlying BiFeO_3 layer²⁴. It also revealed electric-field-driven 90° switching of the local $\text{Co}_{0.9}\text{Fe}_{0.1}$ moments to be a consequence of exchange coupling to the weak ferromagnetism in BiFeO_3 rather than strain-induced effects²⁴ (Methods). Hence, the $\text{Co}_{0.9}\text{Fe}_{0.1}/\text{BiFeO}_3$ structure essentially amplifies the canted moment (6–8 electromagnetic units per cm^3) of BiFeO_3 , making it easily measurable before and after an electric field is applied. The electric-field-driven reversal of D and M_c can thus be experimentally determined with this heterostructure.

$\text{Pt}/\text{Co}_{0.9}\text{Fe}_{0.1}/\text{BiFeO}_3$ and $\text{Pt}/\text{Co}_{0.9}\text{Fe}_{0.1}/\text{Cu}/\text{Co}_{0.9}\text{Fe}_{0.1}/\text{BiFeO}_3$ (spin-valve) heterostructures were patterned into devices to investigate the magnetic state as a function of applied voltage (Fig. 3a). To probe the directions of the moment within each domain and the net moment, X-ray magnetic circular dichroism photoemission electron microscopy (XMCD-PEEM) images of $\text{Pt}/\text{Co}_{0.9}\text{Fe}_{0.1}/\text{BiFeO}_3$ were taken (Methods) in the initial state and after application of 6 V with the incident X-ray aligned perpendicular and parallel to the stripe domains (Fig. 3b, c). After the application of a 6 V (10 ms) pulse, the net $\text{Co}_{0.9}\text{Fe}_{0.1}$ magnetization reversed. Unlike a previous study in which the average magnetization of a large multi-domain structure was reversed as a result of

local 90° rotations using an in-plane electric field (70 V over $6\mu\text{m}$)²⁴, here we demonstrate full 180° electric-field induced reversal of $\text{Co}_{0.9}\text{Fe}_{0.1}$ moments within each domain (Fig. 3d, e), confirming the predicted coupling between the DM vector and ferroelectric polarization.

Anisotropic magnetoresistance measurements were carried out under a 20 Oe magnetic field H before and after the application of ± 7 V (1–10 μs) voltage pulses (applied when $H = 0$) to the $\text{Co}_{0.9}\text{Fe}_{0.1}/\text{BiFeO}_3$ devices depicted in Fig. 3a. Each sequential voltage pulse shifted the anisotropic magnetoresistance curves by about 180° , revealing a reversible switch of the magnetization (Fig. 4a). Two separate devices were switched into the -7 V and 7 V states, similar to Fig. 4a, and the $\text{Co}_{0.9}\text{Fe}_{0.1}$ was removed (Methods) to reveal the BiFeO_3 domains (Fig. 4b). In each case the magnetization reversal is accompanied by a reversal of the polarization, while the orientation of the stripe domain structure is preserved after each switch, mimicking the multi-stage switching process of Fig. 1.

To demonstrate potential technological applicability, we fabricated spin-valve ($\text{Pt}/\text{Co}_{0.9}\text{Fe}_{0.1}/\text{Cu}/\text{Co}_{0.9}\text{Fe}_{0.1}$) devices onto BiFeO_3 films. Figure 4c shows two complete resistance versus voltage, $R(V)$, loops (under zero magnetic field) after the device was configured into a low-resistance, zero-magnetic-field state along with a ferroelectric loop from a neighbouring device. The resistance values are compared to those obtained by resistance versus magnetic field, $R(H)$, curves using dotted lines as visual guides. A clear correlation exists between the ferroelectric switching and the switching from the high- to the low-resistance states of the spin valve. We quantified the energy consumption per switch

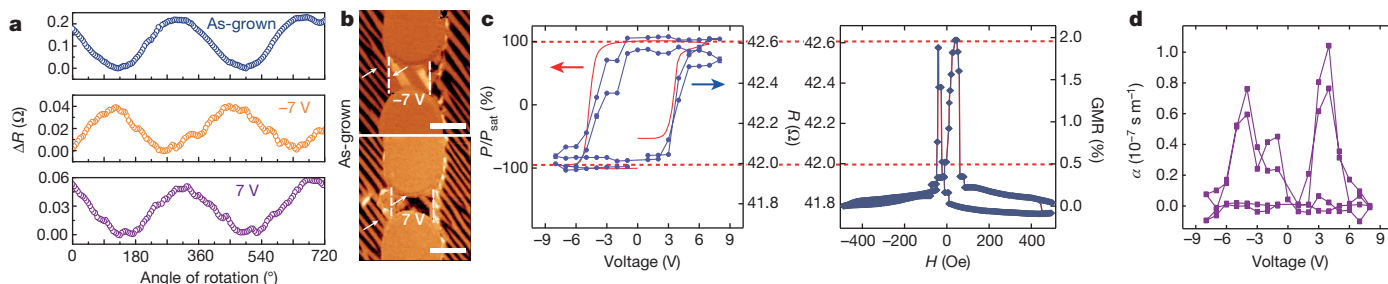


Figure 4 | Magnetoelectric devices. **a**, Anisotropic magnetoresistance versus applied voltage. The upper, middle and lower panels show anisotropic magnetoresistance from the as-grown (as-deposited) state, after -7 V and after 7 V. **b**, In-plane PFM images of two devices in the -7 V (upper) and the 7 V (lower) states after $\text{Co}_{0.9}\text{Fe}_{0.1}$ layers were removed. Scale bars are 2 μm . **c**, The left panel shows two $R(V)$ loops (blue circles) taken under zero magnetic

field plotted along with a ferroelectric loop from a neighbouring device (red line). P_{sat} is the saturation polarization. The right panel shows the $R(H)$ curve correlating $R(V)$ loops to the configuration of the magnetizations in the spin valve and giant magnetoresistance (GMR) values. **d**, Magnetoelectric coefficient α versus applied voltage.

per unit area of our magnetoelectric spin-valve device to be $480 \mu\text{J cm}^{-2}$, roughly an order of magnitude lower than an optimized spin torque device^{13,14} ($3\text{--}4 \text{ mJ cm}^{-2}$; see Methods) and with an applied voltage of only $\sim 4 \text{ V}$. Our results therefore demonstrate that using magnetoelectric multiferroics is a promising strategy for the development of very-low-energy magnetic memory and logic applications at room temperature. We note that these first-generation magnetoelectric devices suffer from reliability issues that probably arise from the metal/oxide interface (Methods).

The $R(V)$ loop can be used to quantify the converse magnetoelectric effect, $\alpha = \mu_0 \frac{dM}{dE}$. Making two simple assumptions about the hysteresis in the observed $R(V)$ curve leads to: $\alpha = \frac{2\mu_0 M_s t}{R_{AP} - R_P} \frac{dR(V)}{dV}$ (Methods), where μ_0 , M_s , t , R_{AP} and R_P are the magnetic permeability of free space, the in-plane saturation magnetization, the thickness of BiFeO_3 and the resistance values of the spin-valve when the magnetizations (in a single-domain/mono-domain states) of the two layers are antiparallel and parallel, respectively. Figure 4d shows that the converse magnetoelectric coefficient reaches giant values ($\alpha \approx 1 \times 10^{-7} \text{ s m}^{-1}$) near 4 V . We estimate, without any further optimization of the magnetoelectric switching, that a single-domain exchange-coupled $\text{Co}_{0.9}\text{Fe}_{0.1}$ layer would increase the value to ($\alpha \approx 3 \times 10^{-7} \text{ s m}^{-1}$). Although giant magnetoelectric coefficients have been reported in single-phase multiferroics^{16,25}, our values are several orders of magnitude larger than those typically observed in single-phase materials^{26,27}. Our values are comparable to those found in ferromagnet–ferroelectric composite structures (composite multiferroics)^{28,29}; nonetheless, they are one to two orders of magnitude smaller than the gigantic value observed in the FeRh/BaTiO_3 system³⁰.

A strength of composite multiferroics (in addition to room-temperature functionality) is a large, typically strain-mediated, magnetoelectric coupling. The symmetry of strain-mediated magnetoelectric coupling, however, precludes deterministic reversal of a magnetization with electric field. In contrast, a composite heterostructure comprised of a room-temperature single-phase multiferroic (in which deterministic switching of its weak ferromagnetism is possible) that is coupled to a strong ferromagnet combines the benefits of single-phase and composite multiferroics. In these heterostructures, deterministic switching of ferromagnetism and a giant converse magnetoelectric effect is achieved using an electric field at room temperature.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 July; accepted 17 October 2014.

- Spaldin, N. A. & Fiebig, M. The renaissance of magnetoelectric multiferroics. *Science* **309**, 391–392 (2005).
- Eerenstein, W., Mathur, N. D. & Scott, J. F. Multiferroic and magnetoelectric materials. *Nature* **442**, 759–765 (2006).
- Fiebig, M. Revival of the magnetoelectric effect. *J. Phys. D* **38**, R123–R152 (2005).
- Scott, J. F. Room-temperature multiferroic magnetoelectrics. *NPG Asia Mater.* **5**, e72 (2013).
- Wang, J. *et al.* Epitaxial BiFeO_3 multiferroic thin film heterostructures. *Science* **299**, 1719–1722 (2003).
- Dzyaloshinskii, I. E. Thermodynamic theory of weak ferromagnetism in antiferromagnetic substances. *Sov. Phys. JETP* **5**, 1259–1262 (1957).
- Moriya, T. Anisotropic superexchange interaction and weak ferromagnetism. *Phys. Rev.* **120**, 91–98 (1960).
- Cheong, S.-W. & Mostovoy, M. Multiferroics: a magnetic twist for ferroelectricity. *Nature Mater.* **6**, 13–20 (2007).
- Sergienko, I. A. & Dagotto, E. Role of Dzyaloshinskii–Moriya interaction in multiferroic perovskites. *Phys. Rev. B* **73**, 094434 (2006).

- Ederer, C. & Fennie, C. J. Electric-field switchable magnetization via the Dzyaloshinskii–Moriya interaction: FeTiO_3 versus BiFeO_3 . *J. Phys. Condens. Matter* **20**, 434219 (2008).
- Ederer, C. & Spaldin, N. A. Weak ferromagnetism and magnetoelectric coupling in bismuth ferrite. *Phys. Rev. B* **71**, 060401(R) (2005).
- Fennie, C. J. Ferroelectrically induced weak ferromagnetism by design. *Phys. Rev. Lett.* **100**, 167203 (2008).
- Liu, H. *et al.* Ultrafast switching in magnetic tunnel junction based orthogonal spin transfer devices. *Appl. Phys. Lett.* **97**, 242510 (2010).
- Rowlands, G. E. *et al.* Deep subnanosecond spin torque switching in magnetic tunnel junctions with combined in-plane and perpendicular polarizers. *Appl. Phys. Lett.* **98**, 102509 (2011).
- Fusil, S., Garcia, V., Barthélémy, A. & Bibes, M. Magnetoelectric devices for spintronics. *Annu. Rev. Mater. Res.* **44**, 91–116 (2014).
- Evans, D. M. *et al.* Magnetic switching of ferroelectric domains at room temperature in multiferroic PZTFT. *Nature Commun.* **4**, 1534 (2013).
- Belik, A. A. *et al.* Neutron powder diffraction study on the crystal and magnetic structures of BiCoO_3 . *Chem. Mater.* **18**, 798 (2006).
- Li, M.-R. *et al.* A polar corundum oxide displaying weak ferromagnetism at room temperature. *J. Am. Chem. Soc.* **134**, 3737–3747 (2012).
- Wang, W. *et al.* Room temperature multiferroic hexagonal LuFeO_3 films. *Phys. Rev. Lett.* **110**, 237601 (2013).
- Benedek, N. A. & Fennie, C. J. Hybrid improper ferroelectricity: a mechanism for controllable polarization–magnetization coupling. *Phys. Rev. Lett.* **106**, 107204 (2011).
- Zanolli, Z., Wojdel, J. C., Íñiguez, J. & Ghosez, P. Electric control of magnetization in $\text{BiFeO}_3/\text{LaFeO}_3$ superlattices. *Phys. Rev. B* **88**, 060102(R) (2013).
- Yang, Y., Íñiguez, J., Mao, A.-J. & Bellaiche, L. Prediction of a novel magnetoelectric switching mechanism in multiferroics. *Phys. Rev. Lett.* **112**, 057202 (2014).
- Lebeugle, D., Mougin, A., Viret, M., Colson, D. & Ranno, L. Electric field switching of the magnetic anisotropy of ferromagnetic layer exchange coupled to the multiferroic compound BiFeO_3 . *Phys. Rev. Lett.* **103**, 257601 (2009).
- Heron, J. T. *et al.* Electric-field-induced magnetization reversal in a ferromagnet–multiferroic heterostructure. *Phys. Rev. Lett.* **107**, 217202 (2011).
- Kimura, T. *et al.* Magnetic control of ferroelectric polarization. *Nature* **426**, 55–58 (2003).
- Rado, G., Ferrari, J. & Maisch, W. Magnetoelectric susceptibility and magnetic symmetry of magnetoelectrically annealed TbPO_4 . *Phys. Rev. B* **29**, 4041–4048 (1984).
- Hur, N. *et al.* Electric polarization reversal and memory in multiferroic material induced by magnetic fields. *Nature* **429**, 392–395 (2004).
- Eerenstein, W., Wiora, M., Prieto, J. L., Scott, J. F. & Mathur, N. D. Giant sharp and persistent converse magnetoelectric effects in multiferroic epitaxial heterostructures. *Nature Mater.* **6**, 348–351 (2007).
- Zhang, S. *et al.* Electric-field control of nonvolatile magnetization in $\text{Co}_{40}\text{Fe}_{40}\text{B}_{20}/\text{Pb}(\text{Mg}_{1/3}\text{Nb}_{2/3})_{0.7}\text{Ti}_{0.3}\text{O}_3$ structure at room temperature. *Phys. Rev. Lett.* **108**, 137203 (2012).
- Cherifi, R. O. *et al.* Electric-field control of magnetic order above room temperature. *Nature Mater.* **13**, 345–351 (2014).

Acknowledgements This work was supported by the National Science Foundation (Nanosystems Engineering Research Center for Translational Applications of Nanoscale Multiferroic Systems) under grant number EEC-1160504 and the DOD-ARO MURI supported by the Army Research Office through agreement number W911NF-08-2-0032 and partially by the FENA-FAME and NSF/MRSEC (DMR-1120296), through the Cornell Center for Materials Research, programs. J.I. acknowledges financial support from MINECO-Spain (grant numbers MAT2010-18113 and CSD2007-00041). J.L.B., L.Y. and B.D.H. acknowledge support from the DOE ESPM (grant number DE-SC0005037). The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under contract number DE-AC02-05CH11231. We thank C. Fennie for discussions, C. Zollner for the figure edits, and H. Nair, G. Stiehl and N. Dawley for their comments on the manuscript.

Author Contributions J.T.H. designed and, with M.T., D.C.R., D.G.S. and R.R., directed the study, analysed results, and wrote the manuscript. J.L.B. and L.Y. performed the time-dependent PFM measurements. J.L.B. and B.D.H. analysed the results. Q.H. and J.L. performed the XMCD measurements and provided analysis. Y.G. and M.T. grew the $\text{BiFeO}_3/\text{SrRuO}_3$ bilayers. J.D.C. aided in device fabrication. C.W. aided in the measurement and analysis of the spin-valve devices. J.I. performed the *ab initio* calculations. S.S. evaluated the transport results. J.T.H., M.T., C.W., S.S., D.C.R., D.G.S., J.L., B.D.H. and R.R. all made contributions to writing the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.T.H. (jth247@cornell.edu).

METHODS

Pulsed-laser deposition of BiFeO₃ (100 nm)/SrRuO₃ (8 nm) bilayers. Films in this study were fabricated by pulsed-laser deposition using a KrF laser on single-crystalline (110) DyScO₃ substrates to obtain a stripe-like, ferroelectric domain structure³¹ composed of two polarization variants. First, a thin layer of conducting SrRuO₃ (8 nm) was deposited and followed by BiFeO₃. Substrate temperatures of 690 °C and 700 °C were used for the deposition of SrRuO₃ and BiFeO₃ films, respectively. The films were grown under oxygen pressures of 50 mTorr and 100 mTorr, respectively, at a repetition rate of 8 Hz with a laser fluence of 1.1 J cm⁻². After growth, the samples were cooled to room temperature in an O₂ pressure of 750 Torr.

Time-dependent PFM. The switching dynamics of the BiFeO₃ (100 nm)/SrRuO₃ (8 nm) structure were measured by time dependent³², dual-frequency³³ PFM using an Asylum Research Cypher in ambient conditions. To activate polarization switching, a 3.515 V direct-current offset was applied to a conductive probe (Nanosensors NCHR) during scanning along the BiFeO₃ [100]_p direction. This was superimposed with a 2 V peak-to-peak amplitude, applied at both the normal cantilever contact resonance frequency of approximately 1.8 MHz, as well as at a lateral resonance of approximately 1 MHz. This generates spectrally distinct piezoresponses along the [001]_p and [100]_p directions that are analysed with a multi-frequency lock-in amplifier (Zurich Instruments HF2LI). The phase signals are used to determine the domain orientation along <111>_p. 85 consecutive frames were recorded to resolve polarization-switching dynamics with a temporal resolution per pixel per image frame of ~40 μs.

Determination of polarization direction in partially switched state. To determine the stable polarization variants accessed during this study, experiments were performed on BiFeO₃ that had been partially switched *in situ*, leaving regions where the multi-step switching was incomplete (Extended Data Fig. 1). PFM images of this area were then obtained before (0°) and after a 90° (in-plane) rotation of the sample. In both cases the normal and one lateral component were measured, such that the <001> as well as the <010> and <100> piezo-vectors are mapped. Each pair of images (Extended Data Fig. 1a, b) reduces the possible orientations from eight to four. Combining these results, the local polarization directions are determined absolutely (Extended Data Fig. 1c). Only four domain (polarization) orientations are observed in the partially switched image.

Ab initio calculations. For the simulations we used the generalized gradient approximation to density functional theory as implemented in the VASP package³⁴. In particular, we used a generalized gradient approximation optimized for solids³⁵. A 'Hubbard-U' scheme with $U = 4$ eV was used for a better treatment of iron's 3d electrons³⁶. We used the 'projector augmented wave' method³⁷ to represent the ionic cores, solving for the 3p, 3d and 4s electrons of Fe; the 5d, 6s and 6p electrons of Bi; and the 2s and 2p electrons of O. We performed our calculations using a 40-atom cell that can be viewed as a 2 × 2 × 2 repetition of the elemental perovskite cell; a 2 × 2 × 2 k-point grid was used for integrations within the Brillouin zone corresponding to this cell. Wavefunctions were represented in a plane-wave basis truncated at 500 eV. These calculation conditions are standard in first-principles studies of BiFeO₃, and were checked to render converged results³⁸. The minimum-energy transition paths were determined using the 'nudged elastic band' method³⁹ implemented in VASP, taking advantage of the extensions provided by the Henkelman group at the University of Texas at Austin⁴⁰. As is customarily done, the epitaxial constraint was imposed by forcing the in-plane lattice vectors of the BiFeO₃ cell to match those corresponding to the (110) DyScO₃ substrate.

Calculations of the switching path. *Ab initio* calculations of the BiFeO₃ system were performed with the in-plane lattice constants constrained to that of an (110)-oriented DyScO₃ substrate, without the consideration of a multidomain structure or an influence from an unswitched matrix.

As previously mentioned, we determined the switching paths at the first-principles level in an automatic way, using the 'nudged elastic band' method to find the most energetically favourable way to transit between the initial and final states that are given as an input of the calculation³⁹. When using the 'nudged elastic band' method, one has to work with a certain number of intermediate configurations that will adapt automatically to find the minimum energy path. In our case, paths were computed between all relevant states (that is, between all symmetry-inequivalent choices of polarization and rotation axis), and we typically used between four and nine intermediate configurations depending on the complexity of the path (for example, we used fewer intermediate points for paths involving the switch of only one polarization component, and more points for paths involving the switch of two or three components).

The lowest-energy switching path calculated for a direct (one-step) 180° switch is shown in Extended Data Fig. 3a. A large energy of ~240 meV per formula unit separates the reversal of the polarization, monitored here by the shift of the Bi³⁺ ion. Along this direct path, the bismuth distortions along the three primary coordinates change simultaneously (that is, along the [111]_p direction; all three components behave nearly identically), passing through a state of zero polarization at

which the off-centring vanishes. Throughout the switching of the bismuth distortion, the O₆-octahedral rotation remains relatively unperturbed and ends in its original position as the polarization reaches its final state.

Surprisingly, the calculations reveal that the lowest-energy switching path occurs in a three-step sequence of 71° switches (where the polarization remains non-zero throughout) which also cause the O₆-octahedral rotation to parallel the changes made by the Bi³⁺ displacement (Extended Data Fig. 3b). This result holds irrespective of whether we perform the calculation with or without the epitaxial strain from the DyScO₃ substrate imposed. It is interesting to note that the two-step switching observed experimentally is not the lowest-energy path calculated. As discussed in the next section, the likely causes of this discrepancy are strain and electrostatic conditions that were not included in our calculations. Note that our *ab initio* simulations have only considered a single ferroelectric domain under the epitaxial strain of the substrate and cannot consider the full experimental conditions. Conditions such as the influence of the striped domain structure of the film (Fig. 1a) and the elastic and electric coupling between these ferroelectric domains or, given that only a small area of the film is switched, the elastic and electrostatic interactions between the switched domains and the unswitched BiFeO₃ matrix at the periphery of this switched region. To obtain the two-step switching from the three-step switching, the out-of-plane 71° switch was not allowed in the calculation. This constraint is in agreement with experimental observation, because the out-of-plane 71° switch was not observed. In fact, of the four possible polarization orientations lying outside the (011)_p plane, none were observed throughout the switching. It is, however, possible that such domain orientations are metastable and switch too quickly to be detected by the PFM experiments performed here. This makes the in-plane 71°, the out-of-plane 109° and 180° the only stable switching events possible. As the in-plane 71° and out-of-plane 109° switches (Fig. 1d) are observed magnetoelectric switches^{23,24,41,42}, a reversal of the DM vector through sequential in-plane 71° and out-of-plane 109° switches would be expected to reverse the weak ferromagnetism.

We note that the Fe³⁺ and Bi³⁺ displacements are essentially identical throughout the pathways studied. Extended Data Fig. 4 shows that the only notable difference between the two displacements is that the Fe³⁺ distortions are smaller in magnitude.

Influence of multidomain and surrounding unswitched BiFeO₃ matrix. The experimentally observed switching has several peculiar features. First, throughout the entire switching sequence, [111]_p (dark blue) domains only touch [111]_p (orange) and [111]_p (light blue) domains; the other domain variants behave analogously. Extended Data Fig. 5 shows a frame in the intermediate stage of the switching process (in a different region of the sample with respect to the images shown in Fig. 1), where this is easily observed. If we assume that our domain walls satisfy the electrostatic condition $\nabla \cdot \mathbf{D} = 0$, then all observed boundaries are compatible with domain walls lying only in a (011)_p plane. This suggests that the combination of electrostatic and elastic conditions strongly influence the permissible domain wall planes in this system.

Second, no out-of-plane 71° switch was observed by PFM. This is interesting because the electric field is applied along [001]_p or [001]_p directions and the out-of-plane 71° switch would naturally seem to be the lowest-energy switch. This observation indicates that elastic constraints on the thermodynamically stable polarization directions strongly influence the switching path in our anisotropically strained films. In fact, of the eight possible polarization directions allowed in rhombohedral BiFeO₃, only four polarization directions are observed throughout the switching and all lie in the (011)_p plane. This suggests that the elastic conditions prevent four of the polarization directions from being occupied (that is, two of the four rhombohedral distortion axes—those contained in the (011)_p plane), which would explain the absence of the expected out-of-plane 71° switch when applying an electric bias along the [001]_p direction.

Third, the domain walls remain coherent across the boundary between switched and as-grown regions (Extended Data Fig. 6). This highlights the system's preference to preserve the (011)_p domain wall and suggests that the surrounding unswitched BiFeO₃ imposes an additional elastic constraint⁴³. Given that some of the polarization domains cannot meet at a domain wall, it seems natural that the unswitched region acts as a nucleation site for the formation and propagation of domain walls in the switched region and can fix the width of the domain after switching. As the unswitched matrix is unaltered by switching of the neighbouring region, the boundary conditions imposed by it remain fixed. This latter point may drive the deterministic nature of this observed polarization switching.

Coupling mechanism leading to Co_{0.9}Fe_{0.1} reversal. In a previous work²⁴ we observed an electric-field-induced magnetization switching of a BiFeO₃/Co_{0.9}Fe_{0.1} heterostructure using an in-plane electric field. The Co_{0.9}Fe_{0.1} moments were observed to switch by 90° in this case after the polarization underwent an in-plane 71° switch (this projects as a 90° switch onto the (001) surface). As the in-plane 71° switch is both magnetoelectric and ferroelastic, a second heterostructure was grown

to test whether the observed switching is driven by exchange coupling or strain coupling. This structure consisted of a thin epitaxial layer of insulating and non-magnetic SrTiO₃ (1 nm) directly on top of the BiFeO₃ to break the exchange coupling with the ferromagnetic Co_{0.9}Fe_{0.1}, but that allows any strain from the ferroelastic switching of the polarization to transfer to the Co_{0.9}Fe_{0.1}. The SrTiO₃ layer in this sample does indeed break the exchange coupling (Extended Data Fig. 7) and electric-field-induced switching of the Co_{0.9}Fe_{0.1} moments was not observed (Extended Data Fig. 8). These data indicate that the mechanism of switching is driven by the switching of the canted moment in BiFeO₃. We note that the Co_{0.9}Fe_{0.1} composition was intentionally chosen for these studies to reduce the influence of magnetoelastic effects as the magnetostriction coefficient goes to zero near this composition⁴⁴.

Last, the magnetic state or magnetic anisotropy of transition metal ferromagnets can be controlled through the modulation of interface charge^{45–47}. These voltage-induced effects in transition metals are typically due to a surface magnetic anisotropy that is dependent on the density of carriers (or filling of electronic states) at the interface. As the out-of-plane component of the polarization switches in our Co_{0.9}Fe_{0.1}/BiFeO₃ heterostructure, modulation of the electronic state of the interface may change upon polarization switching, leading to an alteration of the magnetic state. In such cases, however, the transition metal ferromagnet is restricted to thicknesses near or below 1 nm as voltage (interface charge)-dependent surface magnetic anisotropy must dominate over the volume magnetic anisotropy^{46,47}. The Co_{0.9}Fe_{0.1} layer thickness is 2.5 nm in the Co_{0.9}Fe_{0.1}/BiFeO₃ considered here, so the switching of the magnetization via interface charge modulation is unlikely.

Details of Co_{0.9}Fe_{0.1} and magnetic multilayer growth. After the growth of the BiFeO₃ films, they were immediately inserted into a vacuum sputtering chamber with a base pressure of $\sim 3 \times 10^{-8}$ Torr. Here, Co_{0.9}Fe_{0.1} (2–3 nm) layers and spin-valve devices in the sequence of Co_{0.9}Fe_{0.1} (2.3 nm)/Cu (4–7 nm)/Co_{0.9}Fe_{0.1} (2.5 nm)/Pt (2.5 nm) were deposited by direct-current magnetron sputtering onto the BiFeO₃ films at room temperature in a argon background of 8×10^{-4} Torr. The Co_{0.9}Fe_{0.1} layers were deposited under a 200 Oe growth field to induce a magnetic easy axis in the free layer that is parallel to the magnetic easy axis of the pinned layer in contact with BiFeO₃. The devices were then capped with Pt (2.5 nm) to prevent oxidation of the other layers. When necessary, Co_{0.9}Fe_{0.1}/Pt layers were removed via Ar-ion milling.

XMCD-PEEM imaging of the Co_{0.9}Fe_{0.1} magnetization. XMCD-PEEM measurements were completed at PEEM 3 at the Lawrence Berkeley National Laboratory, Advanced Light Source. Focused X-rays were incident on the sample at an angle of 30° from the surface and formed a spot of $\sim 9 \mu\text{m}$ diameter on the sample. Imaging was done by tuning the photon energy to the Co *L*-edge (780 eV) and the use of right- and left-handed circularly polarized radiation enabled imaging of the ferromagnetic Co_{0.9}Fe_{0.1} domain structure by exploiting the XMCD effect at the Co *L*₃- and *L*₂-edges.

Reliability of devices. We note that it has been a challenge to obtain more than three switching cycles in our devices. We anticipate that the mechanism of failure is due to the motion of ionic species under the large electric field, which could trigger the oxidation of the ferromagnetic metal at the interface of the heterostructure⁴⁸. Irreversible oxidation of the ferromagnet would break the interface exchange coupling in the Co_{0.9}Fe_{0.1}/BiFeO₃ system. Since oxide electrodes are traditionally employed to eliminate such trapping of ionic species at metal-oxide interfaces⁴⁹, this suggests that devices could be made robust by using an oxide as the ferromagnetic layer. Indeed, La_{0.7}Sr_{0.3}MnO₃/BiFeO₃ devices have been switched 11 times without any report or indication of failure⁵⁰; however, owing to differences in exchange coupling⁵¹, room-temperature functionality in that system has yet to be demonstrated.

Comparison of switching energy to spin-transfer torque devices. The motivation for integrating multiferroics into spintronics research has been to find low-energy solutions to spintronics applications^{15,52,53}. We quantify the energy consumption of our device and compare our energy loss to that of modern spin-transfer-torque-driven reversal in state-of-the-art conventional magnetic tunnel junctions^{13,14}. At this time, state-of-the-art spin-transfer-torque devices require a voltage pulse of several hundred millivolts (0.7 V) and 500 ps (ref. 13) and 120 ps (ref. 14) in duration through a 60–70 nm \times 180 nm device, producing an energy dissipation per unit area of 3–4 mJ cm^{−2}. For our multiferroic devices, the energy lost in the device is (4 V)(120 μC cm^{−2}) = 480 μJ cm^{−2}, nearly an order of magnitude lower than a well-optimized spin-transfer-torque device. The energy efficiency demonstrated in our first-generation devices is competitive with the improvements projected by using the spin Hall effect to apply a spin-transfer torque in a three-terminal magnetic tunnel junction⁵⁴. Additionally, we note that several pathways for further optimization of the energy consumption are available, such as A-site doping to reduce the ferroelectric switching voltage and polarization of BiFeO₃ (ref. 55).

Determination of converse magnetoelectric coefficient from *R(V)* loop. The converse magnetoelectric coefficient α can be determined from the *R(V)* loop and the *R(H)* of the spin-valve device, because the resistance of the device is a function of the angle between the magnetization of the pinned (the Co_{0.9}Fe_{0.1} layer in contact

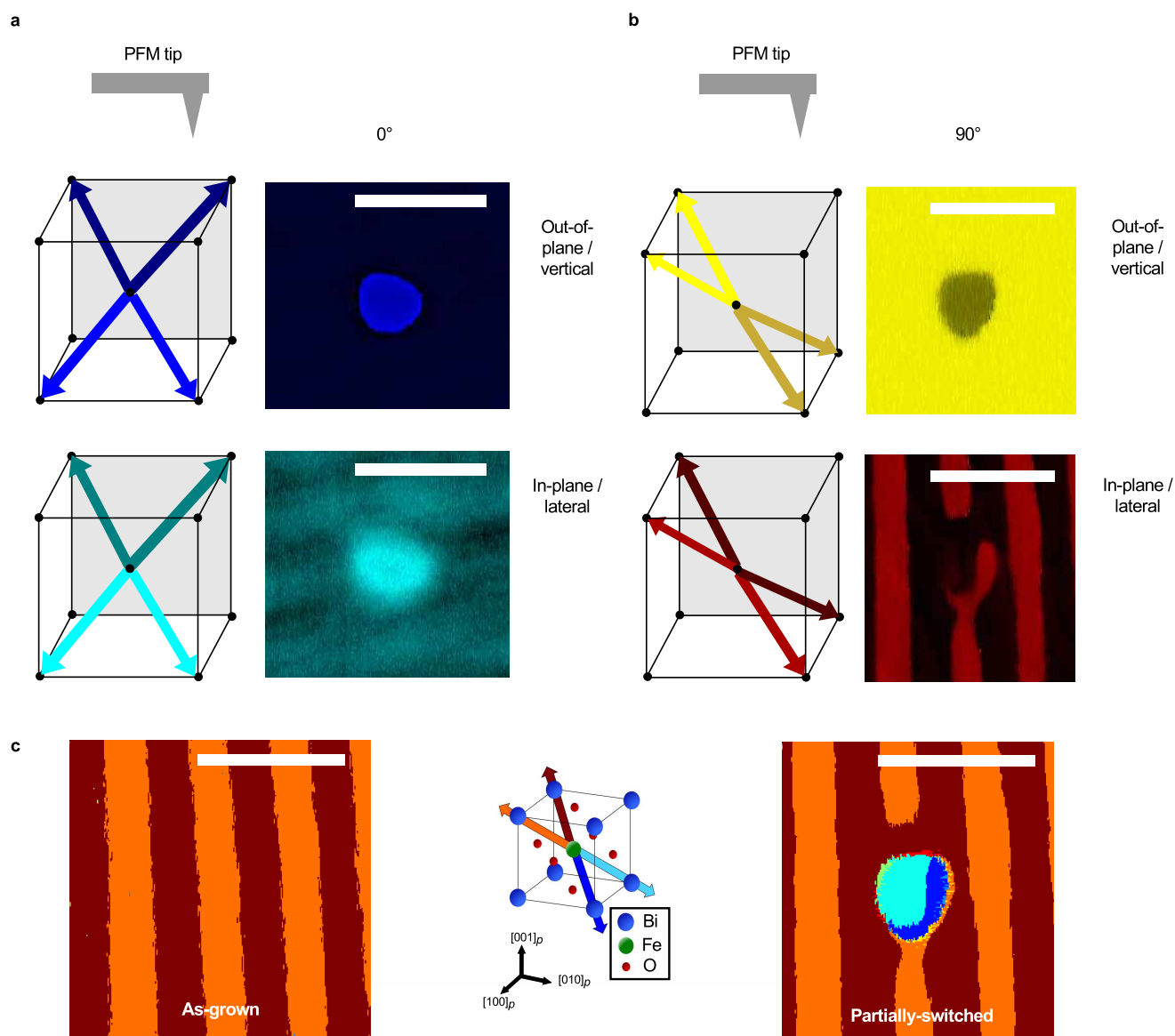
with the BiFeO₃) and free layers (the Co_{0.9}Fe_{0.1} layer not in contact with the BiFeO₃), as follows:

$$R(V, H) = R_p + \frac{(R_{AP} - R_p)(1 - \cos \beta)}{2} \quad (1)$$

where β is the angle between them. It is important to note that because of the coupling between the pinned layer and BiFeO₃ and the devices presented herein being several micrometres in dimension, the pinned layer will tend to adopt the striped-like domain structure of the BiFeO₃ layer^{24,56} (Fig. 3), forming a ‘zigzag’ pattern between the two magnetization directions. Thus the *R(H)* curve will not experience a state where the two magnetizations are both single domain and antiparallel. Extended Data Fig. 9 illustrates the domain structures of the two layers as the field sweeps from positive to negative field (purple curve). In Extended Data Fig. 9 ‘pinned’ and ‘free’ refer to the Co_{0.9}Fe_{0.1} layers in contact and not in contact with BiFeO₃, respectively. So to use equation (1), *R*_{AP} must be determined by other means, whereas *R*_p can easily be obtained from the high-magnetic-field data. If it is then assumed that the free layer is single domain and stays fixed after electrical switching of the pinned layer, then the magnetization of the free layer becomes the reference and the angles between the two magnetizations differ by 180° from the reference after electrical switching. Thus, we can use the high- and low-resistance states of the *R(V)* loop in Fig. 4c and the relationship $\cos(\beta(R_{\text{low}})) = -\cos(\beta(R_{\text{high}}))$ and equation (1) to obtain $R_{AP} = R_{\text{low}} + R_{\text{high}} - R_p$. As the converse magnetoelectric coefficient is given

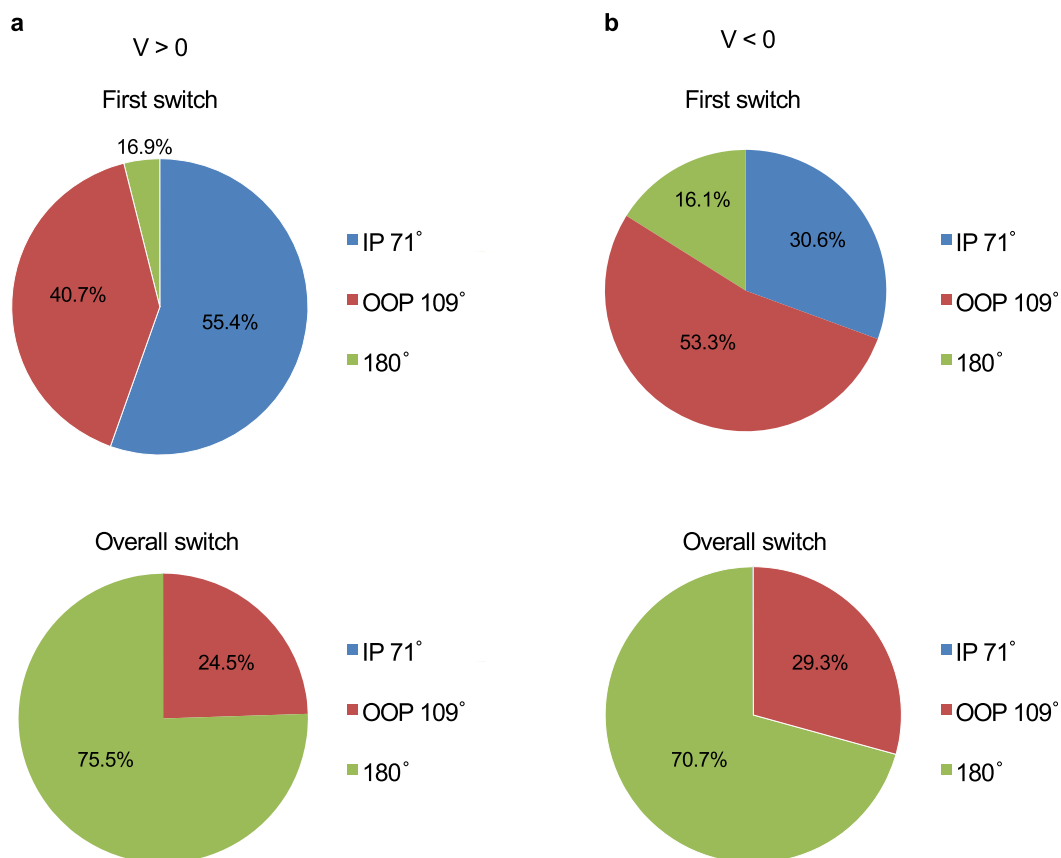
by $\alpha = \mu_0 \frac{dM}{dE}$, where *M* in the reference frame of the magnetization of the free layer is given by $M = M_s \cos \beta = M_s \left(1 - 2 \left[\frac{R(V) - R_p}{R_{AP} - R_p} \right] \right)$, the converse magnetoelectric coefficient is given by $\alpha = \frac{2\mu_0 M_s t}{R_{AP} - R_p} \frac{dR(V)}{dV}$.

- Chu, Y. H. *et al.* Nanoscale control of domain architectures in BiFeO₃ thin films. *Nano Lett.* **9**, 1726–1730 (2009).
- Huey, B. D., Nath, R., Lee, S. & Polomoff, N. A. High speed SPM applied for direct nanoscale mapping of the influence of defects on ferroelectric switching dynamics. *J. Am. Ceram. Soc.* **95**, 1147–1162 (2012).
- Desmarais, J. *et al.* Mapping and statistics of ferroelectric domain boundary angles and types. *Appl. Phys. Lett.* **99**, 162902 (2011).
- Kresse, G. & Furthmüller, J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane wave basis set. *Phys. Rev. B* **54**, 11169 (1996).
- Perdew, J. P. *et al.* Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys. Rev. Lett.* **100**, 136406 (2008).
- Dudarev, S. L., Botton, G. A., Savrasov, S. Y., Humphreys, C. J. & Sutton, A. P. Electron-energy-loss spectra and the structural stability of nickel oxide: an LSDA+U study. *Phys. Rev. B* **57**, 1505 (1998).
- Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953 (1994).
- Diéguez, O., González-Vázquez, O. E., Wojdeł, J. C. & Íñiguez, J. First-principles predictions of low-energy phases of multiferroic BiFeO₃. *Phys. Rev. B* **83**, 094105 (2011).
- Mills, G., Jonsson, H. & Schenter, G. K. Reversible work transition state theory: application to dissociative adsorption of hydrogen. *Surf. Sci.* **324**, 305 (1995).
- Sheppard, D., Xiao, P., Chemelewski, W., Johnson, D. D. & Henkelman, G. A generalized solid-state nudged elastic band method. *J. Chem. Phys.* **136**, 074103 (2012).
- Chu, Y. H. *et al.* Electric-field control of local ferromagnetism using a magnetoelectric multiferroic. *Nature Mater.* **7**, 478–482 (2008).
- Zhao, T. *et al.* Electrical control of antiferromagnetic domains in multiferroic BiFeO₃ films at room temperature. *Nature Mater.* **5**, 823–829 (2006).
- Baek, S. H. *et al.* Ferroelastic switching for nanoscale non-volatile magnetoelectric devices. *Nature Mater.* **9**, 309–314 (2010).
- Hunter, D. *et al.* Giant magnetostriction in annealed Co_{1-x}Fe_x thin-films. *Nature Commun.* **2**, 518 (2011).
- Radaelli, G. *et al.* Electric control of magnetism at the Fe/BaTiO₃ interface. *Nature Commun.* **5**, 3404 (2014).
- Maruyama, T. *et al.* Large voltage-induced magnetic anisotropy change in a few atomic layers of iron. *Nature Nanotechnol.* **4**, 158–161 (2009).
- Wang, W.-G., Li, M., Hageman, S. & Chien, C. L. Electric-field-assisted switching in magnetic tunnel junctions. *Nature Mater.* **11**, 64–68 (2011).
- Couet, S. *et al.* Electric field-induced oxidation of ferromagnetic/ferroelectric interfaces. *Adv. Funct. Mater.* **24**, 71 (2014).
- Ramesh, R. *et al.* Fatigue and retention in ferroelectric Y-Ba-Cu-O/Pb-Zr-Ti-O/Y-Ba-Cu-O heterostructures. *Appl. Phys. Lett.* **61**, 1537 (1992).
- Wu, S. M. *et al.* Reversible electric control of exchange bias in a multiferroic field-effect device. *Nature Mater.* **9**, 756–761 (2010).
- Yu, P. *et al.* Interface ferromagnetism and orbital reconstruction in heterostructures. *Phys. Rev. Lett.* **105**, 027201 (2010).
- Binek, C. & Doudin, B. Magnetoelectronics with magnetoelectrics. *J. Phys. Condens. Matter* **17**, L39–L44 (2005).
- Béa, H., Gajek, M., Bibes, M. & Barthélémy, A. Spintronics with multiferroics. *J. Phys. Condens. Matter* **20**, 434221 (2008).
- Liu, L. *et al.* Spin-torque switching with the giant spin Hall effect of Ta. *Science* **336**, 555 (2012).
- Chu, Y. H. *et al.* Low voltage performance of epitaxial BiFeO₃ films on Si substrates through La substitution. *Appl. Phys. Lett.* **92**, 102909 (2008).
- Trassin, M. *et al.* Interfacial coupling in multiferroic/ferromagnet heterostructures. *Phys. Rev. B* **87**, 134426 (2013).



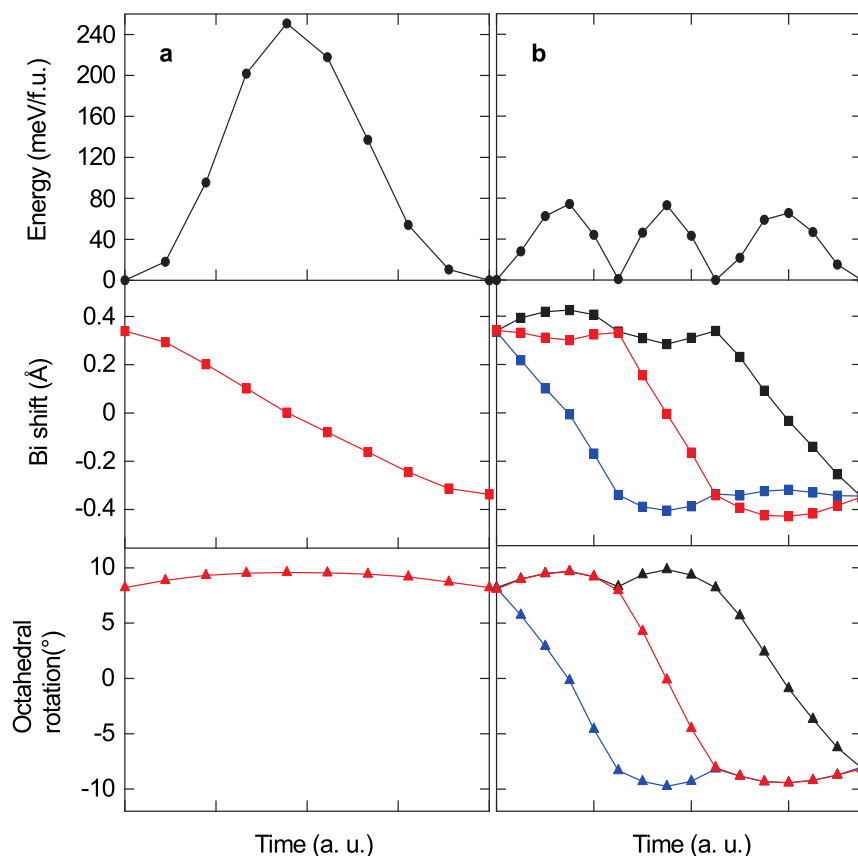
Extended Data Figure 1 | PFM rotational dependence after partial switching. Schematics of the PFM tip and the possible polarization components measured for 0° (a) and 90° (b) (that is, the [010] and [100] directions, respectively), as well as the corresponding vertical and lateral PFM

images upon partial switching. **c**, Schematic of polarization directions and vector PFM images taken before (left) and after partial switching (right). All scale bars are 400 nm.



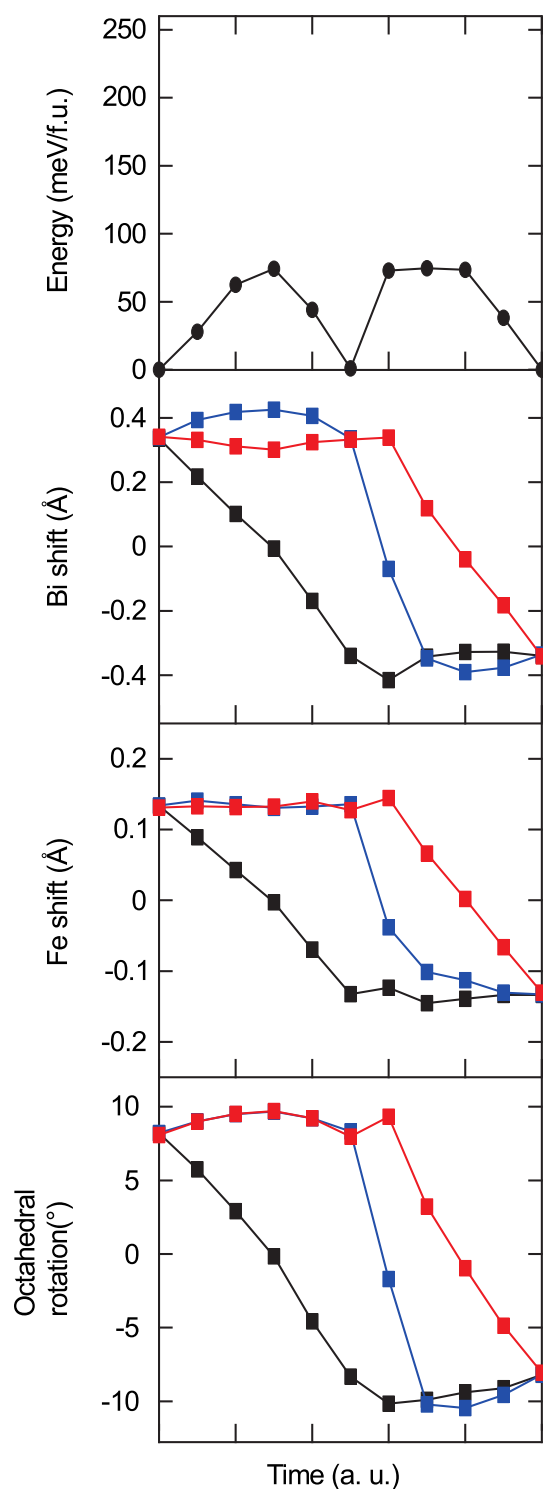
Extended Data Figure 2 | Compositions of the first and overall switching directions under positive and negative PFM tip bias. **a**, Compositions under positive tip bias, which switches the initial or as-grown state. **b**, Under

negative PFM tip bias, which switches the final state of **a** back into an initial configuration. OOP, out of plane; IP, in plane.

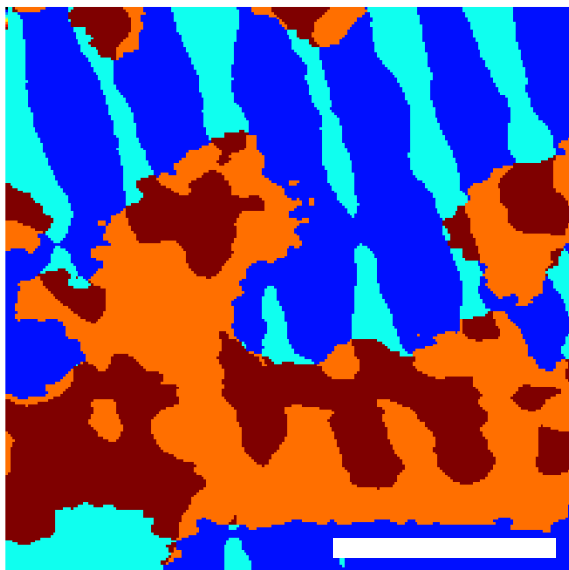


Extended Data Figure 3 | *Ab initio* calculations of the BiFeO₃ system under the constraint of a (110) DyScO₃ substrate. No consideration of the multidomain structure or the influence from an unswitched BiFeO₃ matrix is given. **a**, The lowest-energy switching path calculated for a direct (one-step) 180° switch. A large energy separates the direct reversal of the polarization (black curve, top panel), described here as the Bi³⁺ shift (middle panel). The Bi³⁺ shift reverses following a trajectory directly through zero shift (that is, zero polarization). The O₆-octahedral rotation (lower panel) remains

unperturbed by the direct switch. Black, blue and red curves indicate orthogonal components (x_1 , x_2 , and x_3) of Bi³⁺ displacement and O₆-octahedral rotation in the reference cell. **b**, The lowest-energy switching path for polarization reversal calculated from all possible switching paths is a three-step sequence of sequential ferroelastic 71° switches. In this case the shift of the O₆-octahedral rotation parallels the changes made by the polarization (Bi³⁺ displacement), leading to the reversal of the polarization, octahedral rotation and thus, the weak ferromagnetic moment of BiFeO₃.

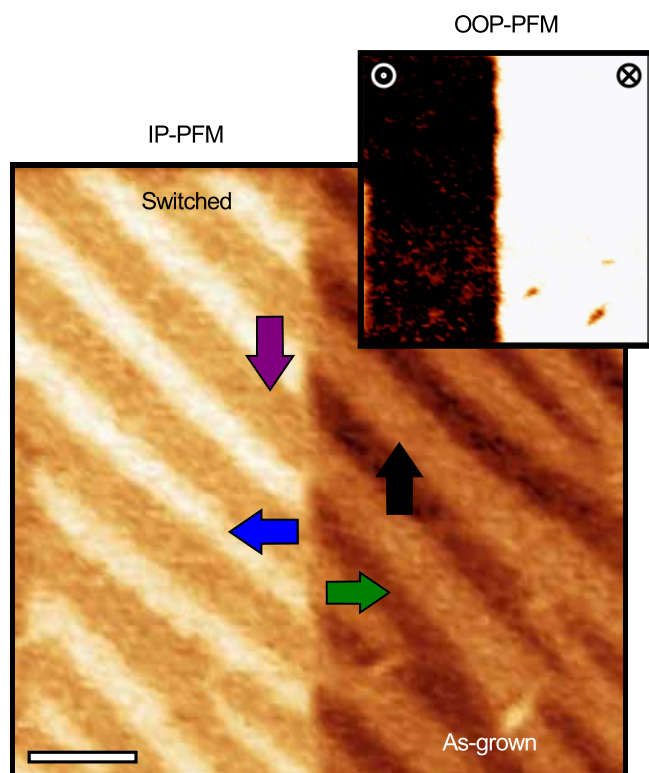


Extended Data Figure 4 | *Ab initio* calculations of the Fe displacement in single- and two-step switching events. Fe^{3+} displacements for the single-step and two-step switching events shown in, and plotted with the data from, Fig. 2a. In each case, the Fe^{3+} shift mimics the Bi^{3+} shift; however, the Fe^{3+} shifts are smaller than those for Bi^{3+} . a.u., arbitrary units. f.u., formula units.

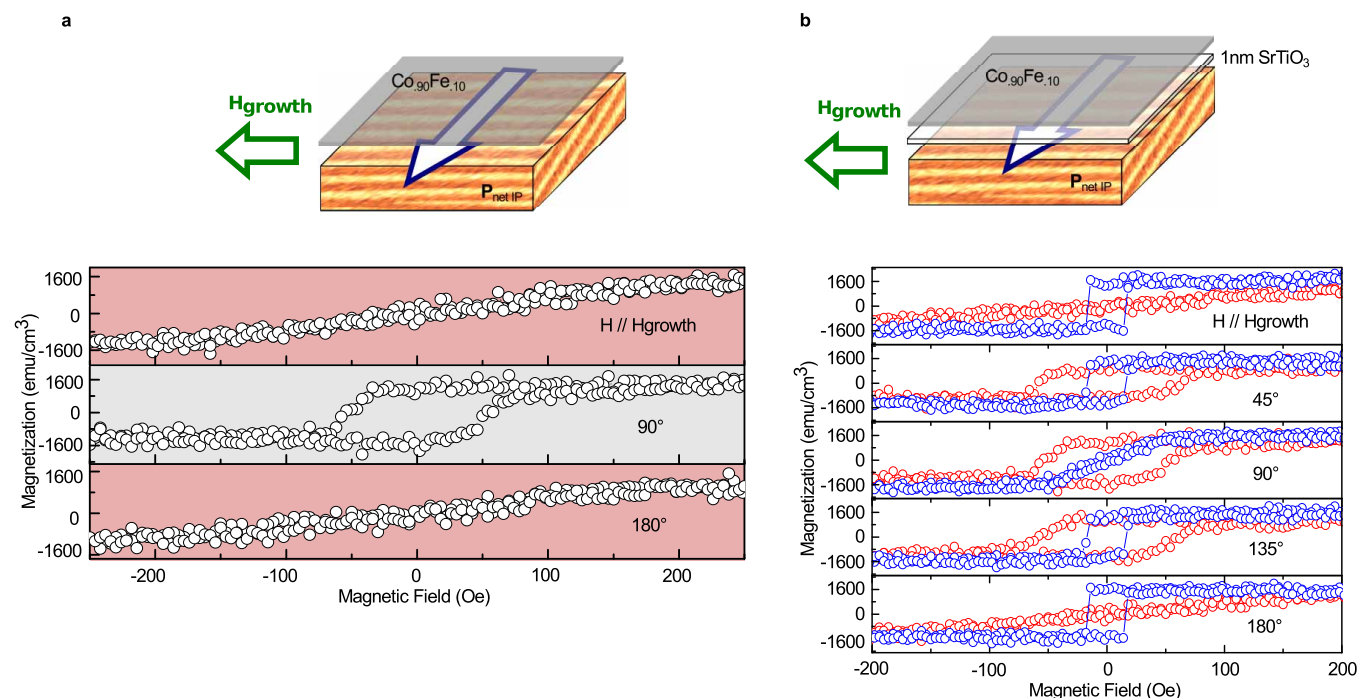
**Extended Data Figure 5 | Domain configuration during switching process.**

A vector PFM image obtained partially through the switching process.

Throughout the entire switching process, and in this image, $[1\bar{1}1]_p$ (dark blue) domains only touch $[1\bar{1}1]_p$ (orange) and $[\bar{1}1\bar{1}]_p$ (light blue) domains; the other polarization directions behave analogously. Scale bar is 500 nm.

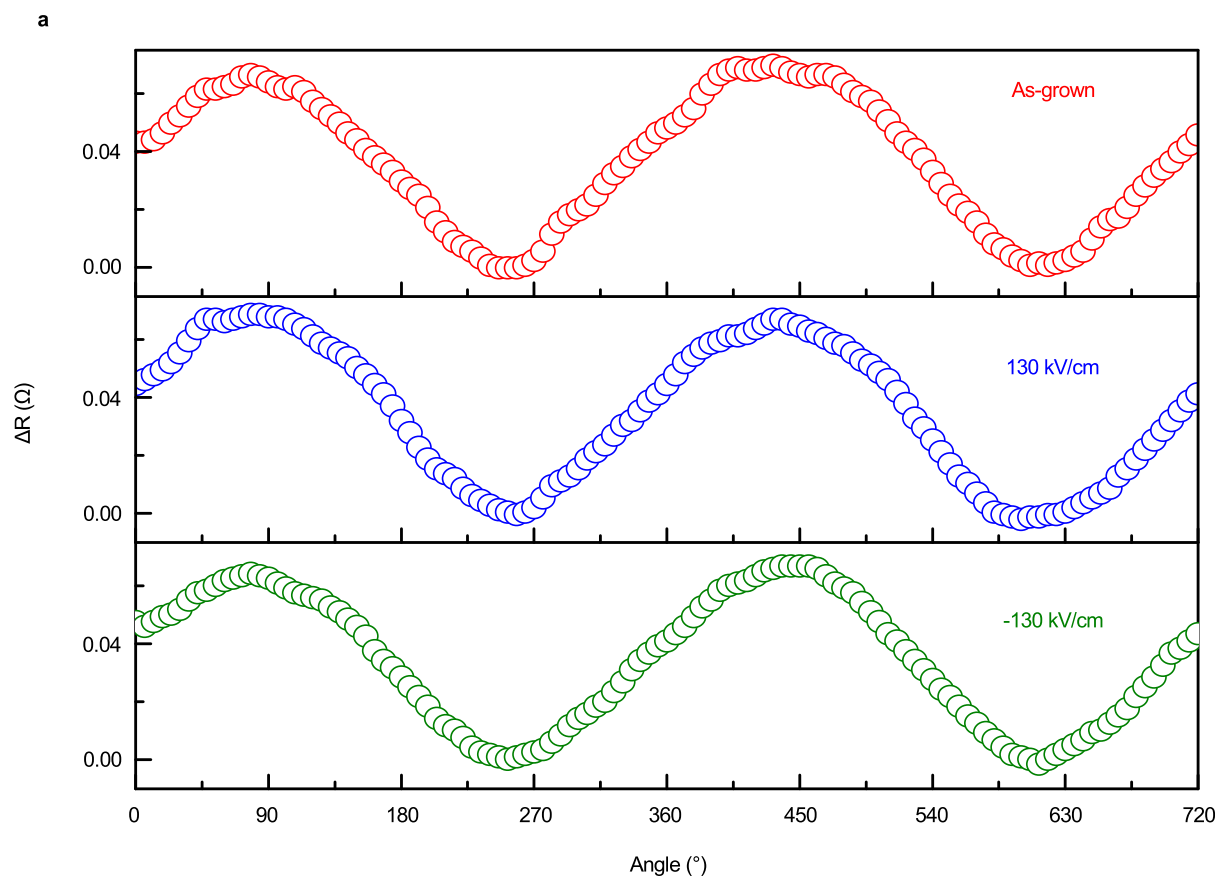


Extended Data Figure 6 | In-plane and out-of-plane PFM images of the boundary between switched and as-grown (initial) regions. The arrows indicate the in-plane and out-of-plane (inset) components of the polarizations. The domain walls across the boundary appear to be continuous, suggesting that the unswitched matrix (as-grown region) has an influence on the final polarization and domain states. Scale bar is 500 nm.



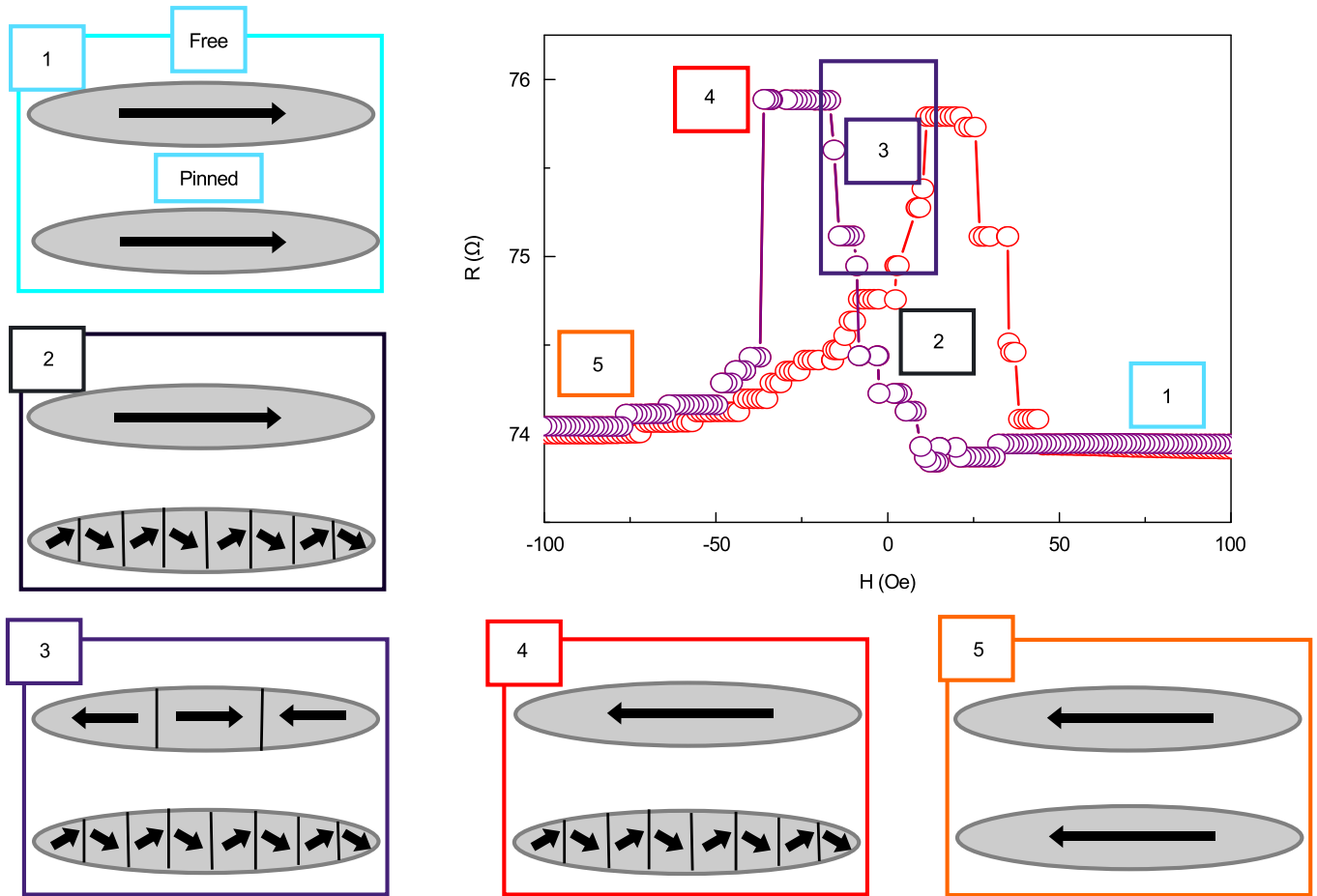
Extended Data Figure 7 | Magnetic hysteresis curves from BiFeO₃/Co_{0.9}Fe_{0.1} (2.5 nm) and BiFeO₃/SrTiO₃ (1 nm)/Co_{0.9}Fe_{0.1} (2.5 nm) heterostructures. **a**, Schematic of the BiFeO₃/Co_{0.9}Fe_{0.1} heterostructure with the directions of the net in-plane polarization (P_{netIP} , the vector sum of the (001)_p surface projections of the two polarization variants) and a 200 Oe magnetic field (H_{growth}) applied during the deposition of the Co_{0.9}Fe_{0.1}. The growth field was used to test the strength of the coupling to BiFeO₃, given that the growth field should attempt to induce a uniaxial anisotropy in that direction. Magnetic hysteresis loops taken along different in-plane angles

rotating from the direction of the growth field show that despite the growth field the easy axis is found to parallel the axis set by P_{netIP} . **b**, Schematic of an experimental configuration similar to that in **a**; however, a 1-nm-thick layer of insulating and non-magnetic SrTiO₃ has been deposited onto the BiFeO₃ before the deposition of Co_{0.9}Fe_{0.1}. The magnetic loops from the BiFeO₃/SrTiO₃/Co_{0.9}Fe_{0.1} heterostructure are plotted in blue and have a uniaxial anisotropy in the direction of the growth field (orthogonal to the axis set by P_{netIP}) with reduced strength (lower saturation and switching fields) compared to those obtained from the BiFeO₃/Co_{0.9}Fe_{0.1} heterostructure in **a** (red curves).



Extended Data Figure 8 | Null electric-field control of magnetism measurement on a $\text{BiFeO}_3/\text{SrTiO}_3$ (1 nm)/ $\text{Co}_{0.9}\text{Fe}_{0.1}$ (2.5 nm) heterostructure. **a**, Anisotropic magnetoresistance obtained from the $\text{BiFeO}_3/\text{SrTiO}_3/\text{Co}_{0.9}\text{Fe}_{0.1}$ heterostructure taken at 20 Oe after the corresponding

electric field was applied. No change in the phase of the anisotropic magnetoresistance curves was observed, indicating no switching of the magnetization.



Extended Data Figure 9 | Description of the magnetoresistance in relation to the domain structure of the unpinned (not in contact with BiFeO_3) and pinned (in contact with BiFeO_3) $\text{Co}_{0.9}\text{Fe}_{0.1}$ layers. As the magnetic field is swept from positive to negative field (open purple circles) along the easy axis of the device the domain structure of the pinned layer evolves from single-domain to a stripe-like structure and back to single-domain. The numbers correlate to the schematics of the domain structures to the spin-valve resistance. At large, positive magnetic field the free and pinned layers are monodomain with magnetizations parallel (1, light blue box) and the device resistance is low. At low, positive magnetic field the pinned layer breaks up into two domain variants owing to the exchange coupling with BiFeO_3 while the free layer

remains largely monodomain (2, black box). Both net magnetizations are parallel but the device resistance increases due to domain formation in the pinned layer. The purple box (3) encloses the region of magnetic field where the unpinned layer breaks up into domains during switching and the device resistance increases rapidly. In box 4 (red) the net magnetizations of the two layers are antiparallel but not fully antiparallel as the pinned layer is broken into domains and the device resistance is high. At high, negative magnetic field the device is again in a low-resistance state and the two layers are monodomain with parallel magnetization. A similar evolution of the domain structure occurs as the magnetic field is increased from negative to positive values (open red circles).

Reconstruction and control of a time-dependent two-electron wave packet

Christian Ott^{1,†}, Andreas Kaldun¹, Luca Argenti², Philipp Raith¹, Kristina Meyer¹, Martin Laux¹, Yizhu Zhang¹, Alexander Blättermann¹, Steffen Hagstötz¹, Thomas Ding¹, Robert Heck¹, Javier Madroño^{3,†}, Fernando Martín^{2,4} & Thomas Pfeifer^{1,5}

The concerted motion of two or more bound electrons governs atomic¹ and molecular^{2,3} non-equilibrium processes including chemical reactions, and hence there is much interest in developing a detailed understanding of such electron dynamics in the quantum regime. However, there is no exact solution for the quantum three-body problem, and as a result even the minimal system of two active electrons and a nucleus is analytically intractable⁴. This makes experimental measurements of the dynamics of two bound and correlated electrons, as found in the helium atom, an attractive prospect. However, although the motion of single active electrons and holes has been observed with attosecond time resolution^{5–7}, comparable experiments on two-electron motion have so far remained out of reach. Here we show that a correlated two-electron wave packet can be reconstructed from a 1.2-femtosecond quantum beat among low-lying doubly excited states in helium. The beat appears in attosecond transient-absorption spectra^{5,7–9} measured with unprecedentedly high spectral resolution and in the presence of an intensity-tunable visible laser field. We tune the coupling^{10–12} between the two low-lying quantum states by adjusting the visible laser intensity, and use the Fano resonance as a

phase-sensitive quantum interferometer¹³ to achieve coherent control of the two correlated electrons. Given the excellent agreement with large-scale quantum-mechanical calculations for the helium atom, we anticipate that multidimensional spectroscopy experiments of the type we report here will provide benchmark data for testing fundamental few-body quantum dynamics theory in more complex systems. They might also provide a route to the site-specific measurement and control of metastable electronic transition states that are at the heart of fundamental chemical reactions.

Electrons are bound to atoms and molecules by the Coulomb force of the nuclei. Moving between atoms, they form the basis of the molecular bond. The same Coulomb force, however, acts repulsively between the electrons. This electron–electron interaction represents a major challenge in the understanding and modelling of atomic and molecular states, their structure and in particular their dynamics^{2,3,14}. Here we focus on the $^1P\ sp_{2,n+}$ series¹⁵ of doubly excited states in helium below the $N = 2$ ionization threshold. They are produced through a single-photon-induced transition of both electrons of the $^1S\ 1s^2$ ground state to at least principal quantum number $n = 2$, and autoionize as a result of electron–electron

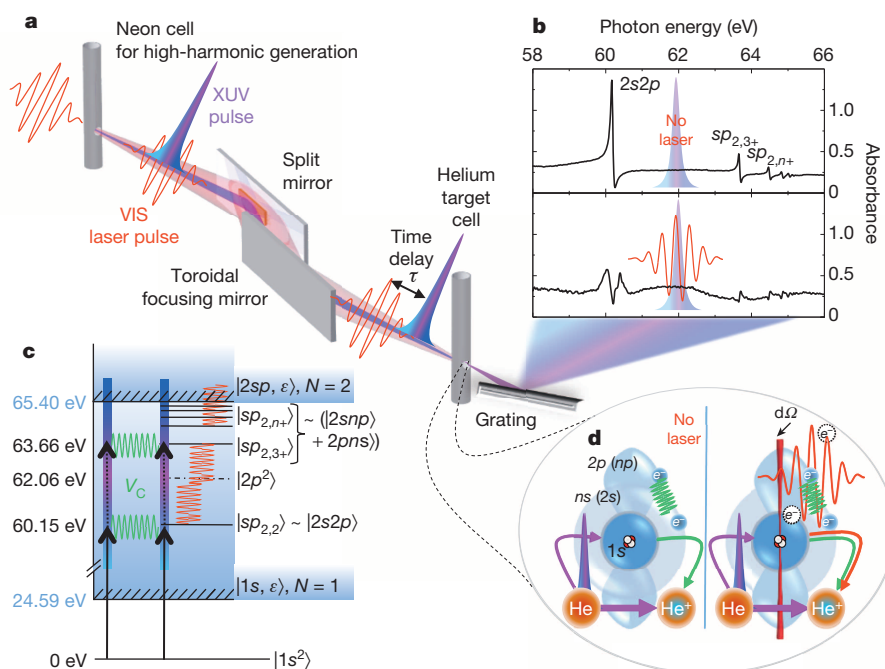


Figure 1 | Experimental set-up, data and microscopic mechanisms in helium. **a**, Few-cycle (7 fs) VIS laser pulses (730 nm) are focused into a neon gas cell for partial attosecond-pulse conversion, providing a continuous coherent excitation spectrum throughout the XUV range. The time delay between the co-propagating VIS and XUV pulses is controlled by a split-mirror stage. Both pulses transit the helium gas target and enter the high-resolution spectrometer. **b**, Absorption spectra without (upper) and in the presence of (lower) the VIS laser pulse, in the region of the $|sp_{2,n+}\rangle$ doubly excited states. **c**, Helium level diagram. The $|sp_{2,n+}\rangle$ states couple (indicated by green wavy lines) to the $|1s, \epsilon p\rangle$ continuum by configuration interaction V_{CI} . The VIS laser field (red wavy lines) creates an additional time- and intensity-dependent coupling. **d**, The XUV pulses can either directly ionize He to He^+ , or excite both electrons into an intermediate transition state, which decays by configuration interaction V_{CI} into He^+ , quantum-interfering with the direct ionization process (left; natural process). If a laser field is present (right), it shifts the phase of one arm of this natural interferometer—the two-electron transition state—modifying the Fano line shapes detected in the transmitted absorption spectrum. This provides state-resolved experimental access to a quantum phase shift.

¹Max-Planck-Institut für Kernphysik, Saupfercheckweg 1, 69117 Heidelberg, Germany. ²Departamento de Química, Módulo 13, Universidad Autónoma de Madrid, 28049 Madrid, Spain. ³Physik-Department, Technische Universität München, 85747 Garching, Germany. ⁴Instituto Madrileño de Estudios Avanzados en Nanociencia, Cantoblanco, 28049 Madrid, Spain. ⁵Center for Quantum Dynamics, Ruprecht-Karls-Universität Heidelberg, 69120 Heidelberg, Germany. [†]Present address: Chemistry Department, University of California, Berkeley, California 94720, USA (C.O.); Departamento de Física, Universidad del Valle, A.A. 25360, Cali, Colombia (J.M.).

interaction. Spectroscopically, the signature of these states is a Fano profile, an asymmetric non-Lorentzian line shape first observed in the 1930s¹⁶ and attributed¹⁷ to the quantum interference of bound states with the continuum to which they are coupled (Fig. 1c, d). The coupling is described by the configuration interaction V_{CI} with the single-ionization continuum $|1s, \epsilon p\rangle$, where one electron is in the $1s$ ground state and the other one is in the continuum with kinetic energy ϵ . The magnitude of V_{CI} determines the lifetimes of the transiently bound states, which in our case range between 17 fs for the $2s2p$ (denoted $sp_{2,2}$) state¹⁸ and several hundreds of femtoseconds for some higher-lying $sp_{2,n+}$ states¹⁵. Such short lifetimes, together with the fast dynamics caused by energy-level spacings on the order of several electronvolts demand ultrashort laser pulses for measuring the coupling dynamics between the states in external fields. Previous time-resolved experiments observed the light-induced modification of absorption profiles⁸, or used attosecond streak-field spectroscopy¹⁸ to measure the $2s2p$ autoionization lifetime. A 1.2 fs two-electron wave packet formed by the coherent superposition of two autoionizing states was recently predicted theoretically¹⁹.

Our experimental method (Fig. 1a, b, ‘Experimental apparatus details’ and ‘Experimental data acquisition’ in Methods, and Extended Data Fig. 1) combines the attosecond transient-absorption scheme and an extreme-ultraviolet (XUV) flat-field grating spectrometer with high-spectral-resolution capability. It allows the parallel measurement of spectrally narrow absorption lines imprinted on an attosecond-pulsed broadband XUV spectrum in the presence of a near-visible (VIS) laser

field. The VIS laser couples the two-electron excited states (Fig. 1c) either weakly, when operated at low intensities, or strongly, when operated at high intensities. The time delay between the VIS and XUV fields and the intensity of the VIS field are varied independently to create a multidimensional transient-coupling scheme that is based on the perturbed free polarization decay and is well known from femtosecond transient-absorption studies²⁰. To complement the experiments, we also performed *ab initio* theoretical calculations of the attosecond transient-absorption spectra and the two-electron wave-packet motion of the helium atom in a laser field using state-of-the-art methods for integrating the time-dependent Schrödinger equation on a fully correlated two-electron close-coupling configuration basis (‘*Ab initio* TDSE simulation’ in Methods).

In Fig. 2, we compare the differential absorption spectra, for varying VIS–XUV time delays and a low VIS intensity of $3 \times 10^{10} \text{ W cm}^{-2}$, obtained from experiment (Fig. 2a), few-level model simulations (Fig. 2b; ‘Few-level model simulation’ in Methods and Extended Data Fig. 2) and *ab initio* calculations (Fig. 2c). We note the excellent agreement between the results, which gives confidence that we can fully understand the dynamics probed in this study. The time-resolved absorption change occurring near the two lowest-lying states, $2s2p$ and $sp_{2,3+}$, appears after zero time delay in the form of temporally oscillating structures with a period of ~ 1.2 fs, indicating coherent two-electron wave-packet dynamics that has been initiated by the XUV pulse and is probed by coupling with the weak VIS pulse. The few-level model confirms the probing mechanism as VIS-induced two-photon dipole coupling of the $2s2p$ and the $sp_{2,3+}$

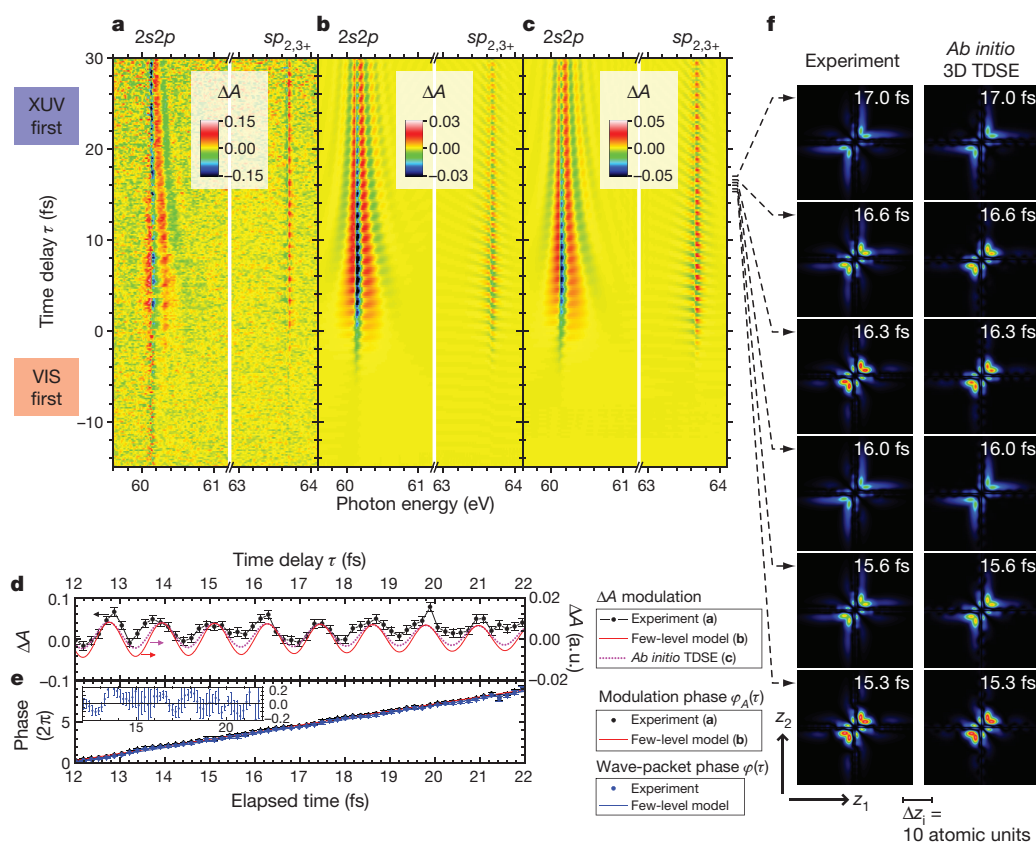


Figure 2 | Observation of attosecond two-electron dynamics in helium. a–c, Absorbance change (ΔA) of XUV light in helium versus time delay between the VIS ($3 \times 10^{10} \text{ W cm}^{-2}$ intensity) coupling field and the XUV pulse: experiment (a), few-level model simulation (b; ΔA in arbitrary units) and *ab initio* calculation (c; ΔA in arbitrary units) show the onset of temporal oscillations near time delay $\tau = 0$ and persisting to large positive delays. d, Oscillation of ΔA (arbitrary units, a.u.) versus τ near resonance at 63.67 eV. e, Modulation phase $\phi_A(\tau)$ of $\Delta A(\tau)$ and relative phase $\phi(t)$ of the XUV-pulse-induced two-electron wave packet involving the $2s2p$ and $sp_{2,3+}$ states, reconstructed by applying to $\phi_A(\tau)$ a small systematic phase shift (‘Measuring the wave-packet phase in real/elapsed time’ in Methods). The inset shows the

experimentally retrieved phase $\phi(t)$ relative to the theoretical expectation. The error bars in d and e reflect the statistical noise (s.d.) of the measured absorption spectra. f, Visualization of the two-electron wave-packet motion. Snapshots of the correlated quantum probability distribution along a line (within infinitesimal cone $d\Omega$, see Fig. 1d) through the helium atom are shown at several instants of elapsed time t . Left column, experimentally reconstructed wave packet including only the two measured states $2s2p$ and $sp_{2,3+}$. Right column, *ab initio* simulation of the three-dimensional time-dependent Schrödinger equation (TDSE), including all excited states. 1 atomic unit = 0.529 Å.

states proceeding via the energetically intermediate and spectroscopically dark $2p^2$ state at 62.06 eV (Fig. 1c). Although the $2s2p \leftrightarrow 2p^2$ transition alone was previously used to control the transmission of helium^{8,12,18}, here we measure and exploit the coupling of three autoionizing states to reconstruct the two-electron wave packet.

The approximate time-dependent wavefunction

$$|\Psi(t)\rangle \propto \exp\left(-\frac{\Gamma_{2s2p}}{2}t\right)|2s2p\rangle + a \exp\left(-\frac{\Gamma_{sp2,3+}}{2}t\right)\exp[-i\varphi(t)]|sp_{2,3+}\rangle$$

is characterized by the relative phase $\varphi(t)$ and amplitude a of the two contributing states, $2s2p$ and $sp_{2,3+}$. The states' slow amplitude decay is given by their respective natural decay widths Γ , accessible from static spectroscopy¹⁵. The relative amplitude $a = (d_{sp2,3+}/d_{2s2p})\sqrt{S(\omega_{sp2,3+})/S(\omega_{2s2p})}$ follows directly from the states' dipole moments d_{2s2p} and $d_{sp2,3+}$ between the doubly excited state and the ground state²¹, and the XUV spectrum $S(\omega)$ at the resonance positions ω_{2s2p} and $\omega_{sp2,3+}$. The relative phase $\varphi(t)$ by contrast is not accessible using traditional spectroscopy. In our time-resolved measurement, different transition pathways involving the doubly excited states interfere as a function of time, allowing us to turn $\varphi(t)$ into an experimental observable by analysing the delay-dependent near-resonance absorption ('Measuring the wave-packet phase in real/elapsed time' in Methods and Extended Data Figs 5 and 6). The measured phase $\varphi(t)$ is plotted in Fig. 2d, e, and agrees well with the *ab initio* simulation results. The relative amplitude is given by $a = 0.5 \pm 0.2$, where the error is mainly due to the fluctuation of the experimental XUV spectrum. The measured values of a and $\varphi(t)$ fully characterize the two-electron wave packet composed of the two autoionizing quantum states $|2s2p\rangle$ and $|sp_{2,3+}\rangle$, which we can reconstruct and visualize by using the known time-independent real-space representations of these states calculated by the complex scaling method²². Figure 2f compares a section of the reconstructed time-dependent spatial distribution of the two electrons against *ab initio* time-dependent simulation, showing very good agreement and that the main features of the two-electron dynamics are thus dominated by the superposition of the $|2s2p\rangle$ and $|sp_{2,3+}\rangle$ states. Owing to the well-defined spectral coherence (that is, phase locking²³) present in a high-harmonic spectrum, the observation of a well-defined phase evolution $\varphi(t)$ is possible²⁴ even in the absence of carrier-envelope phase stabilization and without knowing the number of attosecond pulses in the few-cycle attosecond-pulse train that we generate ('Effects of the attosecond pulse configuration and the carrier envelope phase' in Methods and Extended Data Figs 4–6). We note that the images in Fig. 2f clearly show that the two-electron motion in the reconstructed doubly excited wave packet is highly correlated, although direct experimental observation of such concerted dynamics would require coincidence techniques^{25–27} and represents a major future goal.

The spectra obtained with a higher VIS intensity of $3.5 \times 10^{12} \text{ W cm}^{-2}$ from both experiment (Fig. 3a) and *ab initio* simulations (Fig. 3b) show a shifting, splitting and broadening of the main absorption lines near zero delay, as previously documented in inner-valence excitations of argon⁹. The wave-packet motion is still present and seen as fast absorbance modulations even at late delay times, but is significantly affected by the more intense VIS pulse. Near zero delay time, we also observe strong delay-dependent modifications of the Fano spectral line shapes of the higher-lying states, again with remarkable agreement between the experimental data and the *ab initio* simulation.

After measuring the time-dependent relative phase of quantum states in a two-electron wave packet, we use our experimental method for general two-electron quantum-state holography and wave-packet control. Here the electric field strength of the VIS pulse is an important parameter: it controls the coupling strengths between the states and between the states and the continuum. Continuous variation of the VIS pulse intensity thus opens a third spectroscopic dimension (Supplementary

Video 1), in addition to the time delay and the spectrum. This is illustrated in Fig. 4a, where tuning of the VIS laser intensity at a fixed time delay of 5.4 fs continuously maps the transition from the unperturbed regime to the strong-coupling regime of discrete, doubly excited states that is evident near 60 eV. All states are observed to resist the laser electric fields far beyond classical detachment of the outermost electron by overcoming the attractive nuclear Coulomb force (over-the-barrier ionization²⁸).

We also observe, for all states, a continuous change in the Fano line shape as a function of intensity. Because the line shape of such laser-modified Fano resonances contains information on the phase of the complex dipole response function $d(t)$ after the interaction with the VIS laser pulse¹³, whereas the $1s^2$ ground state is not significantly affected by the VIS laser at the intensities used here, we can use

$$d_n(t) \propto \langle 1s^2 | d | sp_{2,n+} \rangle \exp(-iE_n t) \exp(i\phi_n)$$

to extract the dipole phase shift ϕ_n that is approximately equal to the relative phase shift φ_n of the quantum state at energy E_n . The phase changes of the $2s2p$ and $sp_{2,3+}$ states reconstructed in this way are shown in Fig. 4c, d. The excellent agreement between ϕ_n (extracted from the Fano line shape¹³) and φ_n (defining the wave packet) confirmed by the *ab initio* simulation in Fig. 4b, e, f validates our strategy of experimentally mapping out the intensity-dependent phase shifts $\varphi_n(I)$ of the quantum states by analysing the Fano line shapes ('Line-shape analysis for

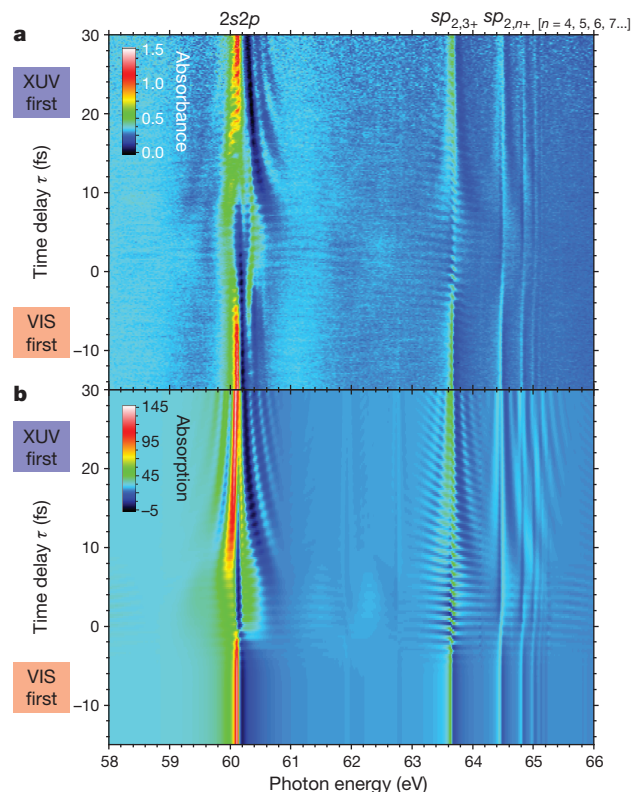


Figure 3 | Time-delay-dependent absorption at higher VIS intensity ($3.5 \times 10^{12} \text{ W cm}^{-2}$). **a**, Time-resolved experimental absorption spectrum at a higher VIS intensity of $3.5 \times 10^{12} \text{ W cm}^{-2}$. At negative delays, the static Fano profile¹⁷ is measured for several autoionizing states up to $sp_{2,7+}$. Near temporal overlap and at positive delays, the absorption spectrum is strongly modified. At slightly positive delays, a clear signature of Autler-Townes splitting of the $2s2p$ resonance with the energetically repelling $2p^2$ dressed state is measured at ~ 60 eV, confirming the strong-coupling regime of autoionizing states and multiple Rabi cycling between these two states. **b**, The full *ab initio* simulation (absorption in arbitrary units) shows excellent agreement with the experiment, thus providing further proof of the existence of a well-defined two-electron wave packet even at higher VIS intensities.

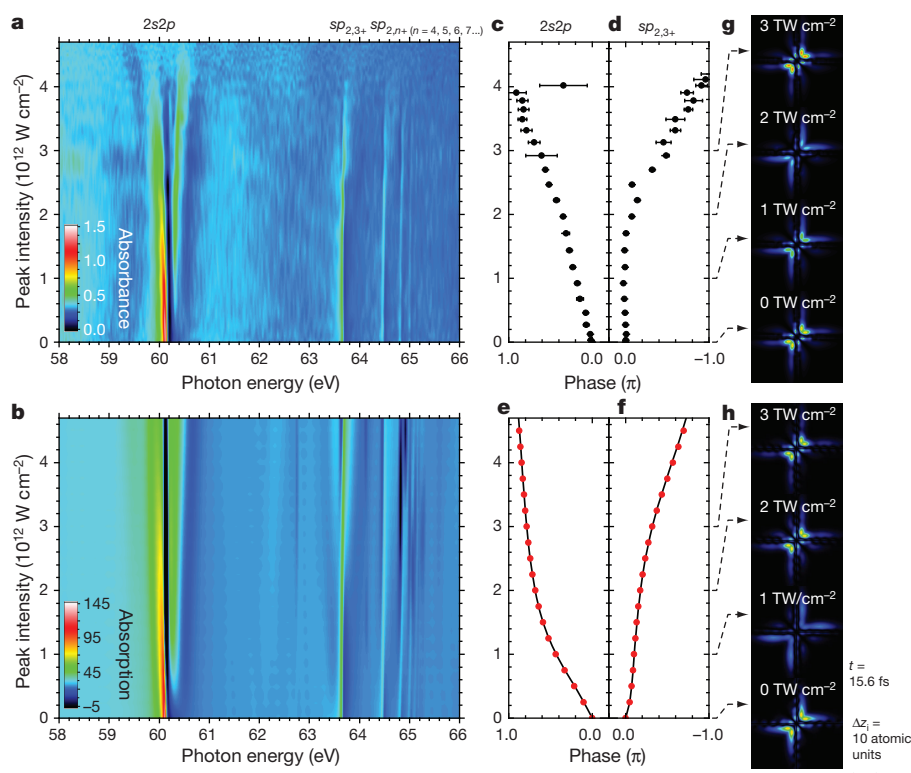


Figure 4 | Intensity-dependent laser coupling and phase control of a two-electron wave packet. **a**, XUV absorption spectra at a time delay of +5.4 fs for increasing VIS coupling intensity. Near 60 eV, we continuously follow the transition from the unperturbed regime to the two-electron strong-coupling regime of the $2s2p$ with $2p^2$ and $sp_{2,3+}$. **b**, Absorption spectra (arbitrary units) calculated using the *ab initio* simulation, as a function of the VIS intensity at a time delay of +4.8 fs. **c**, **d**, Reconstruction of the intensity-dependent temporal phase change of the $2s2p$ (**c**) and $sp_{2,3+}$ (**d**) states after their interaction with the VIS pulse, retrieved via Fano line-shape analysis¹³ ('Line-shape analysis for phase retrieval' in Methods and Extended Data Fig. 7). The state-dependent change of the phase as a function of the intensity

demonstrates laser-controlled manipulation of the two-electron dynamics (**g**), shown in two-dimensional representation at a representative time $t = 15.6$ fs after the XUV pulse. **e**, **f**, Reconstruction of the intensity-dependent phase change in the *ab initio* simulation for the same states as in **c** and **d**, again using Fano line-shape analysis (black solid line). The red dots mark the phase shift of the coefficients, read out after the laser pulse. The excellent agreement with the phase extracted from line-shape analysis (experimentally accessible observable) validates this phase-reconstruction method. **h**, Two-electron probability distribution as obtained from the *ab initio* simulation, for the same time and intensity parameters as in **g**.

phase retrieval' in Methods and Extended Data Fig. 7). Because the phase dependences of $2s2p$ and $sp_{2,3+}$ are opposite in sign, their phase difference can be tuned through $\sim 2\pi$, allowing for the full control of the two-state two-electron wave packet. From the measured laser-induced shift in the phase of each quantum state, we can visualize the shape of the wave packet at any time during its field-free evolution. In Fig. 4g, h, we show this for the representative real time $t = 15.6$ fs. The laser intensity can thus be varied to control the shape of correlated two-electron wave packets at a specific time. In future applications to covalently bound neutral molecules, the strong-field shaping of two-electron wave packets may be a powerful means of laser control of chemical reactions. This further motivates experiments employing coincidence imaging methods (for a review, see, for example, ref. 25) for direct measurements of the spatial shape of two- or multi-electron wave packets as a function of time in the attosecond domain.

The extracted state-dependent phases $\varphi_n(I)$ will give further insight into the coupling between two electrons and how they, collectively or cooperatively, acquire dynamical phases $\varphi_n(I) = \int \Delta E_n(t) dt$ as a result of time-dependent and state (n)-dependent energy-level shifts $\Delta E_n(t)$ (for example Stark or Zeeman shifts) under external perturbation. Having state-resolved access to and control over the full quantum information—amplitude and phase—for two-electron excited states as a function of time and intensity, more fundamental questions can be addressed in the future. For example, how do two-electron transition states respond to field strengths ranging from the weak- to the strong-field limit? What are the dynamics and fate of doubly excited states at and before the onset

of ionization? What is the validity range of commonly used^{18,18,28} single-active electron pictures for strong-field ionization of few-electron systems? The answers have important consequences for goals such as the creation of synthetic atomic quantum systems and exotic molecules using ultrashort, temporally tailored light fields²⁹, beyond the reaches of traditional chemistry.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 August 2013; accepted 28 October 2014.

- Pisharody, S. N. & Jones, R. R. Probing two-electron dynamics of an atom. *Science* **303**, 813–815 (2004).
- Vanroose, W., Martin, F., Rescigno, T. N. & McCurdy, C. W. Complete photo-induced breakup of the H_2 molecule as a probe of molecular electron correlation. *Science* **310**, 1787–1789 (2005).
- Remacle, F. & Levine, R. D. An electronic time scale in chemistry. *Proc. Natl Acad. Sci. USA* **103**, 6793–6798 (2006).
- Lin, C. D. Hyperspherical coordinate approach to atomic and other Coulombic three-body systems. *Phys. Rep.* **257**, 1–83 (1995).
- Goulielmakis, E. *et al.* Real-time observation of valence electron motion. *Nature* **466**, 739–743 (2010).
- Mauritsson, J. *et al.* Attosecond electron spectroscopy using a novel interferometric pump-probe technique. *Phys. Rev. Lett.* **105**, 053001 (2010).
- Holler, M., Schapper, F., Gallmann, L. & Keller, U. Attosecond electron wave-packet interference observed by transient absorption. *Phys. Rev. Lett.* **106**, 123601 (2011).
- Loh, Z. H., Greene, C. H. & Leone, S. R. Femtosecond induced transparency and absorption in the extreme ultraviolet by coherent coupling of the He $2s2p$ ($^1P^o$)

- and $2p^2(^1S^e)$ double excitation states with 800 nm light. *Chem. Phys.* **350**, 7–13 (2008).
9. Wang, H. *et al.* Attosecond time-resolved autoionization of argon. *Phys. Rev. Lett.* **105**, 143002 (2010).
 10. Lambropoulos, P. & Zoller, P. Autoionizing states in strong laser fields. *Phys. Rev. A* **24**, 379–397 (1981).
 11. Themelis, S. I., Lambropoulos, P. & Meyer, M. Ionization dynamics in double resonance involving autoionizing states in helium: the effect of pulse shapes. *J. Phys. B* **37**, 4281–4293 (2004).
 12. Chu, W. C., Zhao, S. F. & Lin, C. D. Laser-assisted-autoionization dynamics of helium resonances with single attosecond pulses. *Phys. Rev. A* **84**, 033426 (2011).
 13. Ott, C. *et al.* Lorentz meets Fano in spectral line shapes: a universal phase and its laser control. *Science* **340**, 716–720 (2013).
 14. Meyer, H. D., Manthe, U. & Cederbaum, L. S. The multi-configurational time-dependent Hartree approach. *Chem. Phys. Lett.* **165**, 73–78 (1990).
 15. Schulz, K. *et al.* Observation of new Rydberg series and resonances in doubly excited helium at ultrahigh resolution. *Phys. Rev. Lett.* **77**, 3086–3089 (1996).
 16. Beutler, H. Über Absorptionsserien von Argon, Krypton und Xenon zu Termen zwischen den beiden Ionisierungsgrenzen $2P_3^{2/0}$ und $2P_1^{2/0}$. *Z. Phys. A* **93**, 177–196 (1935).
 17. Fano, U. Effects of configuration interaction on intensities and phase shifts. *Phys. Rev.* **124**, 1866–1878 (1961).
 18. Gilbertson, S. *et al.* Monitoring and controlling the electron dynamics in helium with isolated attosecond pulses. *Phys. Rev. Lett.* **105**, 263003 (2010).
 19. Argenti, L. & Lindroth, E. Ionization branching ratio control with a resonance attosecond clock. *Phys. Rev. Lett.* **105**, 053002 (2010).
 20. Pollard, W. T. & Mathies, R. A. Analysis of femtosecond dynamic absorption-spectra of nonstationary states. *Annu. Rev. Phys. Chem.* **43**, 497–523 (1992).
 21. Rost, J. M., Schulz, K., Domke, M. & Kaindl, G. Resonance parameters of photo doubly excited helium. *J. Phys. B* **30**, 4663–4694 (1997).
 22. Reinhardt, W. P. Complex coordinates in the theory of atomic and molecular-structure and dynamics. *Annu. Rev. Phys. Chem.* **33**, 223–255 (1982).
 23. Paul, P. M. *et al.* Observation of a train of attosecond pulses from high harmonic generation. *Science* **292**, 1689–1692 (2001).
 24. Kim, K. T. *et al.* Amplitude and phase reconstruction of electron wave packets for probing ultrafast photoionization dynamics. *Phys. Rev. Lett.* **108**, 093001 (2012).
 25. Ullrich, J. *et al.* Recoil-ion and electron momentum spectroscopy: reaction-microscopes. *Rep. Prog. Phys.* **66**, 1463 (2003).
 26. Morishita, T., Watanabe, S. & Lin, C. D. Attosecond light pulses for probing two-electron dynamics of helium in the time domain. *Phys. Rev. Lett.* **98**, 083003 (2007).
 27. Palaudoux, J. *et al.* Multielectron spectroscopy: Auger decays of the krypton 3d hole. *Phys. Rev. A* **82**, 043419 (2010).
 28. Augst, S., Meyerhofer, D. D., Strickland, D. & Chin, S. L. Laser ionization of noble-gases by Coulomb-barrier suppression. *J. Opt. Soc. Am. B* **8**, 858–867 (1991).
 29. Assion, A. *et al.* Control of chemical reactions by feedback-optimized phase-shaped femtosecond laser pulses. *Science* **282**, 919–922 (1998).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank E. Lindroth for calculating the dipole moment $\langle 2p^2 | r | sp_{2,3+} \rangle$, and A. Voitkiv, Z.-H. Loh and R. Moshhammer for discussions. We acknowledge financial support from the Max Planck Research Group Program of the Max Planck Gesellschaft (MPG), the Deutsche Forschungsgemeinschaft (grant no. PF 790/1-1) and the European COST Action CM1204 XLIC. L.A. and F.M. acknowledge computer time from the CCC-UAM and Mare Nostrum supercomputer centres and financial support from the European Research Council under the ERC Advanced Grant no. 290853 XCHEM; the Ministerio de Economía y Competitividad projects FIS2010-15127, FIS2013-42002-R and ERA-Chemistry PIM2010EEC-00751; and the European grant MC-ITN CORINF.

Author Contributions C.O. and T.P. designed the experiment. C.O., A.K., P.R., R.H. and T.P. built the experimental apparatus. C.O., A.K., Y.Z., P.R. and T.P. conducted the experiments. C.O., A.K., S.H., R.H. and T.P. performed the experimental data analysis and interpretation. C.O., A.K., S.H., T.D., R.H. and T.P. carried out the few-level model simulation. L.A. designed and implemented the *ab initio* methods. L.A. and F.M. carried out the *ab initio* simulation and significantly contributed to the interpretation of the results. J.M. calculated the field-free wave functions. C.O., A.K., L.A., F.M. and T.P. wrote major parts of the manuscript. All authors contributed to the discussion of the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials related to the experiment and the few-level model should be addressed to T.P. (thomas.pfeifer@mpi-hd.mpg.de), and those related to the *ab initio* calculations should be addressed to F.M. (fernando.martin@uam.es).

METHODS

Experimental apparatus details. The laser system (commercial Femtolasers compact Pro Ti:Sapphire multipass amplifier; hollow-core fibre spectral broadening; temporal pulse compression with chirped mirrors) typically delivers sub-7 fs, ~ 730 nm, 0.3 mJ laser pulses at 4 kHz repetition rate. The carrier-envelope phase (CEP) was not stabilized but averaged over to avoid additional fluctuations from CEP noise in the high-harmonic spectrum, especially since our measurement scheme is insensitive to the CEP (see ‘Effects of the attosecond pulse configuration and the CEP’ below). The vacuum set-up is shown in Extended Data Fig. 1a. The laser pulses were focused (50 μm focal-spot size; peak intensities, 10^{14} – 10^{15} W cm $^{-2}$) into a stainless-steel cell filled with neon gas, entering and exiting through 100 μm -diameter machine-drilled holes in the cladding. A small fraction of the visible (VIS) light was up-converted into the extreme ultraviolet (XUV) energy range, using high-harmonic generation (HHG) for attosecond pulse production³⁰. The macroscopic parameters were optimized for continuous spectra (100 mbar neon backing pressure; cell position near the laser focus). In Extended Data Fig. 1b, a typical XUV spectrum is shown, alongside an XUV spectrum after transmission through a 100 mbar helium gas target. The co-propagating XUV and VIS pulses were separated by a 2 μm silicon nitride membrane with a ~ 2 mm-diameter centre hole, in combination with a concentrically mounted 200 nm aluminium filter behind the hole. This separation scheme makes use of the intrinsically lower divergence of the XUV beam. The time delay between the XUV and VIS pulses was obtained by a grazing-incidence (15°) split-mirror set-up consisting of an inner gold-coated mirror (2 mm size) for the XUV, and a surrounding silver mirror for the VIS. The inner mirror can be translated with respect to the outer mirror using a high-precision piezoelectric stage (~ 1 nm resolution; ~ 260 μm range). Both beams were refocused (1:1 geometry; 350 mm focal length) with a gold-coated toroidal mirror under the same 15° grazing angle of incidence into another stainless-steel cell filled with helium gas. The monolithic set-up guarantees a high interferometric stability (measured temporal precision, ~ 10 as), combined with broadband and high-throughput advantages of all-grazing-incidence optics. Spectral selection was achieved using thin metal filters (200 nm aluminium), transmitting in the 20–70 eV energy range³¹. The intensity of the VIS beam on the helium target was finely tuned using a picomotor-controlled iris diaphragm centred around both beams. The XUV radiation transmitted through the helium target was spectrally imaged using a flat-field spectrometer consisting of a variable-line-spacing (VLS) grating and a thermoelectrically cooled, back-illuminated XUV CCD camera. The VIS stray light was removed with a pair of 200 nm aluminium metal filters. The spectrometer calibration was obtained by identifying the observed $sp_{2,n+}$ two-electron resonance lines in helium and using tabulated experimental values of high-precision synchrotron measurements^{15,32}. The spectral resolution ($\sigma = 20$ meV Gaussian standard deviation) near 60 eV resulted from a fit of the $2s2p$ resonance line. The target gas density (~ 100 mbar) was chosen such that the strongest $2s2p$ absorption line was still well below absorbance $A = 2$ to avoid dispersion and propagation effects³³. The zero position of time delay was obtained by generating high harmonics in argon in the target gas cell, and accounting for the known thickness of the silicon nitride membrane and the aluminium filter.

Experimental data acquisition. Sets of XUV spectra were recorded as a function of time delay (from -18 to $+34$ fs in ~ 170 as steps; negative values correspond to VIS pulse arriving first) and VIS intensity (35 different iris diaphragm opening settings up to the 10^{12} W cm $^{-2}$ peak-intensity regime), where the intensity calibration was obtained *in situ* as described below (‘Intensity calibration’). Each single spectrum was obtained by integration over $\sim 3,200$ laser shots. For each VIS intensity, additional XUV spectra were recorded without the target helium gas to obtain reference spectra (Extended Data Fig. 1b). The absorbance A is obtained from the general formula $A = -\log_{10}(I_{\text{SIG}}/I_{\text{REF}})$, where I_{SIG} is the signal and I_{REF} is the reference spectral intensity. At an exemplary VIS intensity of 3.3×10^{12} W cm $^{-2}$, the two-dimensional absorbance, plotted versus time delay and photon energy, is displayed in Extended Data Fig. 1c. All relevant structures as discussed in the main text can already be seen. The noisy structures (horizontal lines) are a result of the non-simultaneous measurement of signal and reference XUV spectra, and were filtered out for our quantitative analysis using the following method. For each recorded signal XUV spectrum (containing absorption lines), a low-pass Fourier filter was used to filter out the ‘slowly’ modulating (~ 3.4 eV period) high-harmonic XUV spectrum. This *in situ* filtered spectrum $I_f(\omega)$ was scaled to obtain a reconstructed reference spectrum $I_{\text{REF},\text{rc}}(\omega)$ using Beer’s law: $I_{\text{REF},\text{rc}}(\omega) = I_f(\omega) \exp[\sigma_{\text{PCS}}(\omega)lp]$, where σ_{PCS} is the known non-resonant photo-absorption cross-section of helium³⁴. The path-length/density product lp is the free scaling parameter and was determined to be $lp = (0.56 \pm 0.05) \times 10^{18}$ cm $^{-2}$ via comparison with the measured spectral intensity I_{REF} . As a result of this reference-reconstruction method, the statistical noise of the two-dimensional absorbance plots is significantly reduced, as can be seen by comparing Extended Data Fig. 1c with Fig. 3a. The differential absorption spectra shown in Fig. 2 were generated by subtracting the field-free (no VIS laser interacting

with the XUV-induced dipole at early delays) static spectra, plotting the change of the absorbance (ΔA).

Few-level model simulation. The model system consists of three autoionizing states, $2s2p$ ($^1P^o$), $2p^2$ ($^1S^e$) and $sp_{2,3+}$ ($^1P^o$), at excitation energies 60.15 eV, 62.06 eV and 63.66 eV, respectively, above the $1s^2$ ($^1S^e$) helium ground state^{32,35}. These autoionizing states will be referred to as $|a\rangle$, $|b\rangle$ and $|c\rangle$, respectively, in the following. Other states belonging to the $N = 2$ doubly excited Rydberg series are off-resonance with respect to the coupling VIS laser (~ 1.7 eV photon energy) and/or are significantly lower in coupling strength, and are thus neglected. This subsystem of states is sufficient to reproduce the experimentally observed 1 fs quantum beat. The model is based on previous work^{10,11} solving the time-dependent Schrödinger equation in the configuration basis of the VIS-coupled states. The parity-allowed ($^1S^e \leftrightarrow ^1P^o$) transitions are expressed by the dipole matrix elements d_{nm} as depicted in Extended Data Fig. 2a, which also includes the configuration-interaction matrix elements $V_{e,n}$ that connect the autoionizing states to their respective single-electron continua $|1s, \epsilon s\rangle$ or $|1s, \epsilon p\rangle$. In accordance with earlier approaches for a similar system¹¹, the non-resonant VIS-induced coupling of the $^1P^o$ states with the $|1s, \epsilon s\rangle$ continuum is neglected. Also the VIS coupling between the two continua can be safely neglected in our intensity regime¹⁰. Extended Data Fig. 2b depicts the Schrödinger equation of the so-described few-level system, with the states’ complex expansion coefficients $c_n(t)$, and using atomic units. The weak excitation with the broadband XUV field $F_{\text{XUV}}(t)$ is described in first-order perturbation theory, that is, $\partial_t c_g(t) = 0$, with $E_g = 0$. Under the rotating-wave approximation, $F_{\text{XUV}}(t)$ is taken as a complex quantity, neglecting the anti-resonant part of the interaction. The coupling between the excited bound states is mediated using the full time-dependent real representation of the VIS field $F_{\text{VIS}}(t)$. The continua are treated in the strong-field approximation as Volkov states with the vector potential $A_{\text{VIS}}(t) = -\int_{-\infty}^t dt' F_{\text{VIS}}(t')$ and are parameterized by their canonical momentum p . A one-dimensional treatment is justified owing to the linear polarization of the electric fields. The continuum states are described as quasi-discrete non-interacting states separated in momentum by Δp . To suppress continuum revivals that are an artefact of this discretization, a constant decay rate γ is employed which spectrally broadens the quasi-discrete states to a mutual overlap. The configuration-interaction matrix elements $V_{e,n} = \langle 1s, \epsilon | H | n \rangle \equiv V_n$, which describe autoionization, are taken to be constant (that is, energy independent) in the vicinity of each configuration state, in accordance with Fano’s original theory¹⁷.

Direct numerical integration of the time-dependent complex expansion coefficients $c_n(t)$ was performed with a split-step-like approach, where, for each time interval Δt , different subsystems were evaluated separately. The corresponding five steps were as follows. (i) The perturbative excitation of states $|a\rangle$, $|c\rangle$ and the set of $|1s, \epsilon p\rangle$ continuum states in the XUV laser field. (ii) The coupling of the three bound states $|a\rangle$, $|b\rangle$ and $|c\rangle$ in the VIS laser field. (iii) The coupling of the three bound states $|n\rangle$ with their corresponding continuum states $|1s, \epsilon p\rangle$ and $|1s, \epsilon s\rangle$ owing to configuration interaction. (iv) The field-free evolution of the three bound states with eigen-energies E_n . (v) The VIS-laser-dressed evolution of the quasi-discrete continuum states. For each of these five steps, the corresponding subsystem was diagonalized, and temporal evolution thus corresponds to the multiplication of a complex phase factor $\exp(-i\lambda_j \Delta t)$, with λ_j being the eigenvalues of the diagonalized subsystem after a unitary transformation. For each time point, the time-dependent dipole moment $D(t) = d_{ga}c_a(t) + d_{gb}c_b(t) + d_{gc}c_c(t) + d_g \sum_i c_{\epsilon p,i}(t)$ between the ground state $|g\rangle$ and the dipole-allowed $|a\rangle$ and $|c\rangle$ states as well as the $|1s, \epsilon p\rangle$ continuum states was evaluated, where the ground-continuum dipole matrix element $d_g \equiv d_{gg}$ was assumed to be independent of energy. The absorption spectra were calculated²⁰ via the Fourier transform of $D(t)$, which is proportional to the polarization $P(\omega)$ of the system. Dividing this quantity by the XUV laser spectrum $F(\omega)$ (to obtain a quantity that, in the absence of the VIS field, is proportional to the susceptibility $\chi(\omega)$ of helium, where $P(\omega) = \epsilon_0 \chi(\omega) F(\omega)$; in the presence of the VIS field this corresponds to a generalized linear susceptibility as discussed in ref. 20) and taking the imaginary part of this ratio leads to the XUV absorption profile. This quantity is proportional to the experimentally reconstructed absorbance introduced in ‘Experimental data acquisition’ in our limit of low absorption and, thus, negligible propagation and dispersion effects.

The numerical parameters used (in the respective atomic units) were the discretization time step $\Delta t = 1$ a.u. (0.0242 fs); the total simulation time $T = 32,000$ a.u. (774 fs); the discretized single-electron continuum momentum, which ranged from $p_{\text{min}} = \pm 1.35$ a.u. (that is, $E_{\text{min}} = 24.8$ eV) to $p_{\text{max}} = \pm 2.80$ a.u. (that is, $E_{\text{max}} = 106.7$ eV) in 100 steps with $\Delta p = \pm 0.0145$ a.u. (that is, in total 400 quasi-discrete continuum states) and decay rate $\gamma = 0.1$ a.u.; the energies, widths and asymmetry parameters of the $^1P^o$ states^{15,32} $E_a = 60.147$ eV, $\Gamma_a = 37$ meV and $q_a = -2.75$, and $E_c = 63.658$ eV, $\Gamma_c = 10$ meV and $q_c = -2.53$; the energy and width of the $^1S^e$ state^{35,36} $E_b = 62.06$ eV and $\Gamma_b = 6$ meV; and the dipole matrix element $d_{bb} = 2.17$ a.u., which was taken from ref. 8, whereas $d_{bc} = -0.81$ a.u. was calculated (E. Lindroth, personal communication, 2011). The remaining V_m , d_{gn} and d_g were determined for the simulated absorption spectra to match known experimental and theoretical line shapes with above printed values. The laser pulses were defined as $F^{(0)} \exp[-(t/t_G)^2] \cos(\omega_c t + \varphi_{\text{CEP}})$,

with peak electric field strength $F^{(0)}$; Gaussian pulse duration $t_G = t_p/\sqrt{2\ln(2)}$, where t_p denotes the full-width at half-maximum intensity; the centre frequency ω_c ; and the carrier-envelope phase (CEP) φ_{CEP} . The time discretization and total time simulated allowed us to correctly describe all dynamics in a reasonable amount of computation time (the narrowest linewidth of $F_b = 6$ meV corresponds to a ~ 110 fs lifetime). The decay rate γ effectively maintains the autoionized electrons for ~ 16 a.u., which is a reasonable upper estimate for the spatial extent of the localized two-electron states.

The simulation was validated by using a quasi-monochromatic ($t_G \gg T$) VIS laser field ($\lambda \approx 730$ nm) of increasing field strength $F^{(0)}_{\text{VIS}}$, where the obtained cycle-averaged absorbance is depicted in Extended Data Fig. 2c. As expected, line splitting and a.c. Stark shifts according to the Rabi frequency $d_{nm}F^{(0)}_{\text{VIS}}$ occur owing to Rabi cycling among the three states. To approach the situation realized in the experiments, the temporal evolution of their coefficients is shown in Extended Data Fig. 2d–f at various time delays, where a 7 fs VIS laser pulse was applied instead of the quasi-monochromatic laser field. Significant rearrangement of population between the states occurs, and is maintained after the VIS pulse interaction. This intuitively illustrates how the VIS laser pulse affects the relative population of the two-electron states, which is experimentally accessible in the measured absorption line shapes because these are derived from the oscillating dipole moment $D(t)$. The few-level model simulation was used in the reconstruction of the experimentally observed two-electron wave packet, which allowed for an independent comparison with the full *ab initio* 3D TDSE calculation, and to check for possible effects of various different pulse configurations on the investigated dynamics.

Ab initio TDSE simulation. The *ab initio* transient absorption spectrum was reproduced using the velocity-gauge perturbative expression

$$\sigma_{\text{TAS}}(\omega) = \frac{4\pi}{\omega} \text{Im} \frac{\tilde{p}(\omega)}{\tilde{A}(\omega)} \quad (1)$$

where ω is the field angular frequency, \tilde{p} is the Fourier transform of the total electronic canonical momentum expectation value

$$\tilde{p}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt e^{-i\omega t} p(t)$$

$$p(t) = \langle \psi(t) | p_z | \psi(t) \rangle$$

\tilde{A} is the Fourier transform of the XUV vector potential amplitude, and $\psi(t)$ is the wavefunction for the helium atom in the presence of the external field. The use of equation (1) is justified in the limit of optically thin samples. Already for VIS pulse intensities of the order of few TW cm^{-2} , the optical response $p(t)$ depends non-perturbatively on the VIS external field. For this reason, $p(t)$ was obtained by integrating the TDSE

$$i\partial_t \psi(t) = [H_0 + V_{\text{abs}} + \alpha \tilde{A}(t) \cdot \vec{p}] \psi(t)$$

where H_0 is the field-free electrostatic Hamiltonian of helium, $\alpha \tilde{A}(t) \cdot \vec{p}$ is the minimal-coupling term that accounts for the interaction of the atom with the external field, and V_{abs} is a symmetric complex local potential that prevents reflection from the boundary of the quantization box where the wavefunction is defined. To solve the TDSE accurately, the wavefunction was expanded on the eigenstates of H_0 , projected on a two-particle B-spline close-coupling basis with pseudostates^{37,38}. In such a basis, the angular part is represented by bipolar spherical harmonics and the radial part by B-splines with an asymptotic knot spacing of 0.5 a.u. Each total angular momentum comprises all the partial-wave channels with configurations of the form $Nl\epsilon_r$ with $N \leq 2$, and a full-CI localized channel $nln'l'$ that reproduces short-range correlations between the two electrons. In the presence of the field, the TDSE was integrated numerically with a second-order midpoint exponential time-step propagator¹⁹,

$$\psi(t + dt) = e^{-iV_{\text{abs}} dt} e^{-iH_0 dt/2} e^{-idt\alpha \tilde{A}(t+dt/2) \cdot \vec{p}} e^{-iH_0 dt/2} \psi(t)$$

The action of the second exponential, which depends on the external fields and couples blocks with different symmetry, is evaluated with an iterative Krylov-space method. In the simulation, we included states with total angular momentum up to $L_{\text{max}} = 10$ and, for the localized channel, orbital angular momentum up to $l_{\text{max}} = 5$. We ascertained the convergence of the theoretical results with respect to the most relevant expansion parameters in the state representation by conducting additional representative simulations with either $L_{\text{max}} = 15$ or $l_{\text{max}} = 10$, as well as by including the $N = 3$ partial-wave channels in the close-coupling expansion. After the external field has vanished, the Hamiltonian no longer depends on time and the propagation becomes trivial:

$$\psi(t) = \sum_i |\varphi_{R,i}\rangle e^{-iE_i(t-t')} \langle \varphi_{L,i} | \psi(t') \rangle$$

Here the states $\varphi_{L/R}$ are the left and right eigenstates of the quenched Hamiltonian $H_Q = H_0 + V_{\text{abs}}$ whose complex eigenvalues E_i have non-positive imaginary components:

$$H_Q = \sum_i |\varphi_{R,i}\rangle E_i \langle \varphi_{L,i}|$$

$$\langle \varphi_{L,i} | \varphi_{R,j} \rangle = \delta_{ij}$$

$$\text{Im } E_i \leq 0 \quad \forall i$$

The expectation value $p(t)$ can conveniently be split into two smooth components

$$p(t) = p^-(t) + p^+(t)$$

defined in such a way that $p^-(t)$ becomes negligible shortly after the external field has vanished, whereas $p^+(t)$ is zero before and while the external field is non-zero. As a result of the separation, the Fourier transform of $p(t)$ also splits into the sum of two terms

$$\tilde{p}(\omega) = \tilde{p}^-(\omega) + \tilde{p}^+(\omega)$$

The first Fourier transform was evaluated numerically from tabulated values of $p(t)$ on a dense time grid. The second Fourier transform was instead evaluated analytically from the spectral resolution of the quenched Hamiltonian, the dipole transition matrix elements from the ground state, and the expansion coefficients of the wavefunction at a given time after the external fields have vanished on the numerical basis used to conduct the time propagation. This way of proceeding provides the same result as an infinite time propagation. To compare with the wave packets reconstructed from the experiment, the spatial part of the theoretical wave function $\psi(z_1, z_2; t)$ was tabulated as a function of the Cartesian coordinates z_1 and z_2 when both the electrons are aligned to the field polarization axis, for selected pump–probe time delays and observation times.

Intensity calibration. The simulated few-level dynamics, which is in good qualitative and quantitative agreement with the experimental data and is thus well understood, was used to assess the intensity of the VIS pulse in the interaction region. Both for the numerical and for the experimental results, small temporal regions (averaged over two modulation periods) around 0 fs and around ~ 5 fs time delay (where the Autler–Townes splitting of the $2s2p-2p^2$ doublet is strongest; for the numerical results this was ~ 3 fs) were averaged and the spectra were plotted as a function of the VIS intensity (in the simulation; Extended Data Fig. 3a) and iris diaphragm openings (in the experiment; Extended Data Fig. 3b). By quantifying the induced a.c. Stark shifts of the light-induced $2s2p-2p^2$ Autler–Townes doublet in the experiment, and comparing these shifts with the simulated data (based on the known and experimentally confirmed⁸ dipole matrix element between the two states) an *in situ* intensity calibration was achieved, shown in Extended Data Fig. 3c. The calibration includes an average over various VIS pulse durations (ranging from 5 to 30 fs) to account for the effect of a >7 fs pedestal in the VIS pulse, which is typical of the hollow-fibre/chirped-mirror pulse compression method employed.

Effects of the attosecond pulse configuration and the CEP. The experimental data were obtained by averaging over the CEP. In addition, the coherent XUV excitation spectrum consisted of a train of few attosecond pulses, which is indicated by energy modulations on top of the broad XUV spectra as shown in Extended Data Fig. 1b. Both these effects are negligible for the purposes of observing the discussed effects, as will be shown in the following. Three different XUV pulse configurations have been simulated and are plotted in Extended Data Fig. 4a, and show no significant changes in the absorbance spectra. This is a direct consequence of the well-known²³ phase locking of the attosecond pulses to the half-cycles of the generating intense VIS pulses. In the energy domain, this corresponds to a well-defined coherent excitation spectrum over a broadband spectral range. The insensitivity to the XUV pulse configuration was also confirmed experimentally by performing measurements with and without CEP stabilization. By comparing the corresponding plots in Extended Data Fig. 4b, we see that the additional CEP stabilization in the experiment does not modify the results obtained in the absence of CEP stabilization. Thus, to avoid any sources of error from an imperfect absolute CEP determination (because such a determination does not at present exist for transient-absorption measurements), CEP temporal drift correction, spatial effects such as potential inhomogeneity across the beam profile, or Gouy phase slips in the exact experimental interaction region (extended He target cell), we measured the bulk of our data in the well-reproducible situation of non-stable (and, thus, fully statistical) CEP. We also confirmed the insensitivity of the measurement to the exact pulse-train configuration in the weak-field VIS interaction case for which we extracted the wave-packet phase information shown

in Fig. 2. The results for single-, double- and multi-attosecond-pulse excitation situations are shown in Extended Data Fig. 5. To define the attosecond pulse and its time of arrival with respect to the (generating) VIS laser pulse, we used the coherence (phase locking) between two harmonics, which has been confirmed numerous times to be present in high-harmonic generation since its first direct measurement via interferometric two-photon photoelectron spectroscopy²³. Assuming this phase locking of two harmonics in the energy region 60–64 eV produced the attosecond pulse trains in the model simulation, and also defined the individual attosecond pulse duration to be ~ 600 as.

Measuring the wave-packet phase in real/elapsed time. In the experiment, we measured changes of the spectrally resolved absorbance (ΔA), as a function of time delay τ between an attosecond-pulsed XUV excitation and a VIS coupling pulse. For the case of weak VIS intensity, the coupling process can be considered a weak-field probe process, which does not significantly affect the phases nor the population of the quantum states $2s2p$ and $sp_{2,3+}$ contributing to the wave packet:

$$|\Psi(t)\rangle \propto \exp\left(-\frac{\Gamma_{2s2p}}{2}t\right)|2s2p\rangle + a \exp[-i\varphi(t)] \exp\left(-\frac{\Gamma_{sp_{2,3+}}}{2}t\right)|sp_{2,3+}\rangle$$

In the simulations shown in Extended Data Fig. 5, conducted at the same intensity as for the experimental results in Fig. 2a, we confirmed that population transfer to the near-resonant $2p^2$ state was below 10%. In that weak-field case, the measured ΔA – τ data can be converted into information on the wave-packet states' relative-phase evolution $\varphi(t)$ in real time t (elapsed time after excitation). To define elapsed/real time zero, we used the arrival time of the exciting attosecond pulse, or the most intense central attosecond pulse in the case of a short pulse train, as depicted in Extended Data Fig. 5. A lineout of ΔA versus τ at a spectral position near the $sp_{2,3+}$ resonance at 63.67 eV (where ΔA shows pronounced changes with τ) is shown in Extended Data Fig. 6a and used to map $\Delta A(\tau)$ to $\varphi(t)$. The oscillation of $\Delta A(\tau)$ is almost fully independent of whether the excitation occurs with isolated attosecond pulses or with pulse trains of two or several attosecond pulses. The wave-packet phase $\varphi(t)$, defined as the time-dependent phase difference between the $2s2p$ and $sp_{2,3+}$ state coefficients, was read out from the simulation for all pulse configurations and compared with the phase $\varphi_A(\tau)$ of the oscillation with $\Delta A(\tau) \propto \cos[\varphi_A(\tau)] + \text{const}$ for $t = \tau$, as shown in Extended Data Fig. 6b. The phase $\varphi_A(\tau)$ was retrieved via Fourier analysis, taking the full modulation bandwidth into account as shown in Extended Data Fig. 5g, h. The phases of the wave packet, as excited by the different pulse configurations, are in excellent agreement, again showing that the wave packet is well defined even in the absence of isolated attosecond pulses or CEP locking. The difference between $\varphi_A(\tau)$ (measurable quantity) and $\varphi(t)$ (the relative phase between the quantum states defining the wave packet) was extracted and is shown in Extended Data Fig. 6c. This phase difference is almost independent of the XUV excitation configuration (isolated attosecond pulses versus trains of attosecond pulses), and was thus used in the experiment to retrieve the wave-packet phase as a function of elapsed time t from the measured $\Delta A(\tau)$ data. The experimental result is shown in Fig. 2d, e, where the error bar on the experimental wave-packet phase reconstruction as shown there includes the small error given by the experimental uncertainty in the exact attosecond pulse-train configuration, as discussed here. In Extended Data Fig. 6d, we also show that the amplitude ratio a of the wave packet remains well defined (within 10% amplitude-ratio fluctuations) despite the differences in the XUV excitation configurations. Fluctuations of the order of 10% in the high-harmonic spectra are typically present also in CEP-stabilized laser systems driving HHG, either by CEP noise or shot-to-shot driver-pulse intensity noise.

Line-shape analysis for phase retrieval. As was demonstrated in ref. 13, the Fano q asymmetry parameter can be directly related to a phase shift φ of the temporal dipole response after δ -like excitation via

$$\varphi = 2\arg(q - i) \quad (2)$$

This phase shift can be controlled using a short-pulsed laser field as described in the main text in connection to Fig. 4. The laser-controlled phase manipulation of

the states right after excitation can thus be read out by fitting a Fano line shape to the measured absorption spectrum. The absorption line shape obtained from the *ab initio* simulation (shown in Fig. 4b), is directly fitted with an asymmetric Fano line profile, using

$$S_{\text{FANO}} = \frac{a}{q^2 + 1} \left[\frac{(q + \varepsilon)^2}{1 + \varepsilon^2} - 1 \right] + b \quad (3)$$

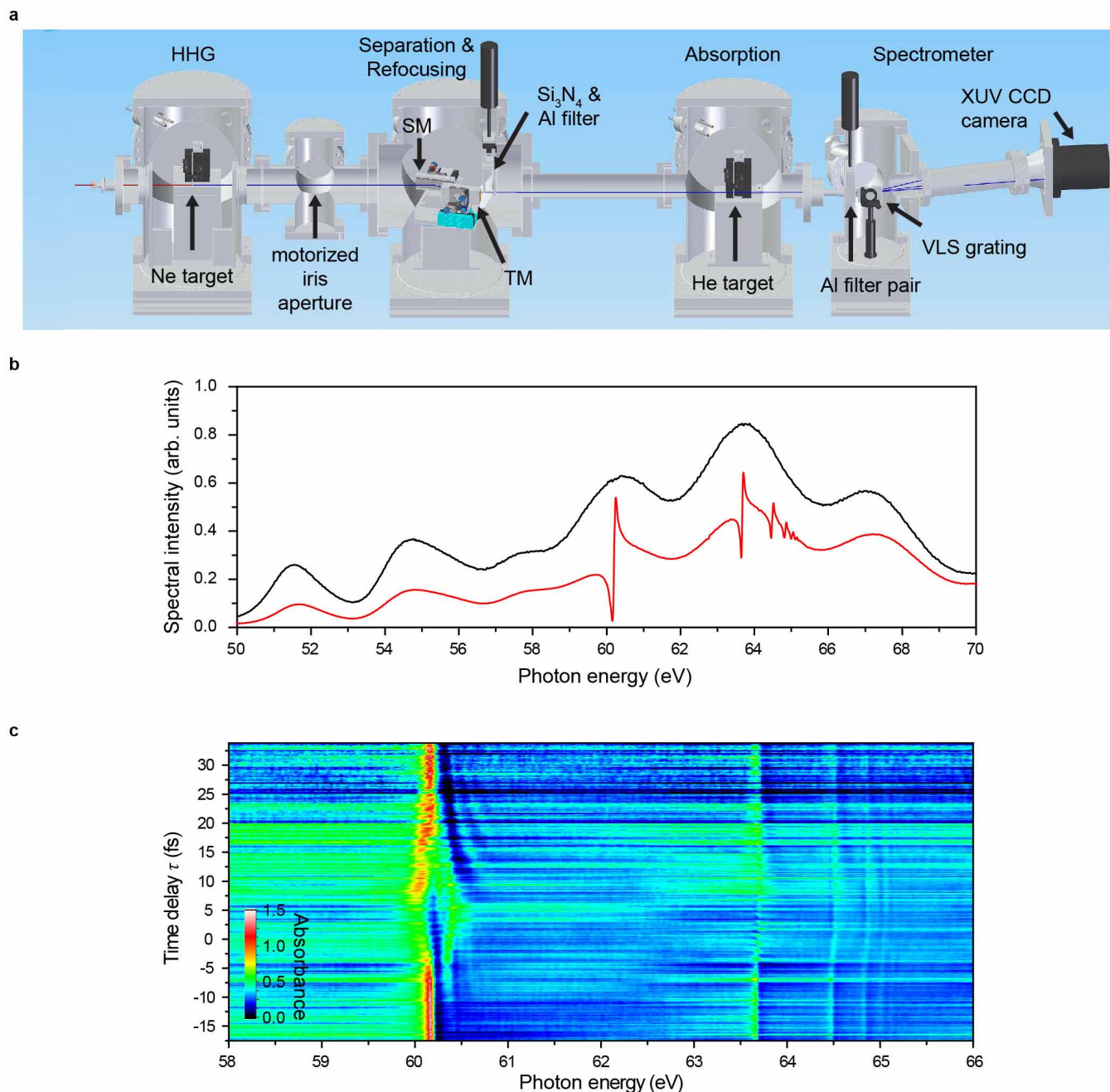
$$\varepsilon = \frac{E - E_r}{\Gamma/2}$$

where ε is the reduced energy. Here all parameters such as the strength a , the offset b , the asymmetry parameter q , the resonance position E_r and the decay width Γ converged to a least-squares minimum. Both the fitted intensity-dependent amplitude $a(I)$ and the phase $\varphi(I)$, where the latter was obtained from q after using equation (2), perfectly agree with the states' complex expansion coefficients. This is shown for the intensity-dependent phase $\varphi(I)$ of the $2s2p$ and $sp_{2,3+}$ states in Fig. 4e, f. Extended Data Fig. 7c, d shows the related fitted line shapes for several VIS laser intensities in the energy region where the least-squares fit was performed, which is 60.11–60.21 eV for the $2s2p$ state and 63.56–63.76 eV for the $sp_{2,3+}$ state. In fitting the experimentally recorded line shapes, we took into account the finite spectrometer resolution, which is of the order of the decay width of the states. Because the XUV intensity $I_{\text{SIG}}(E)$ was measured after transmission through the helium target, $10^{-S_{\text{FANO}}}$ needed to be convolved with the spectrometer response function, which excluded the formulation of an analytical fit function. Formally, the experimentally observed line shape is parameterized via

$$S_{\text{FANO,EXP}} = -\text{Log}_{10} \left[(10^{-S_{\text{FANO}}}) \otimes \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{E^2}{2\sigma^2}\right] \right]$$

where \otimes denotes the convolution, S_{FANO} is given in equation (3) and $\sigma = 0.020$ eV is the experimentally determined detector resolution. In the presence of experimental noise and a limited number of data points, both E_r and Γ were kept constant for all VIS laser intensities using literature values. $S_{\text{FANO,EXP}}$ was numerically computed in the parameter space spanned by q , a and b , and the error sum of mean squared $S_{\text{FANO,EXP}}$ values with respect to the experimental data points was minimized within the same energy region as above. The results are shown in Extended Data Fig. 7a, b and confirm the convergence of the numerically performed minimization procedure. The intensity-dependent phase $\varphi(I)$, extracted from Fano line-shape analysis, is shown in Fig. 4c, d. The error bars were determined by fitting three equivalent experimental data sets and computing the standard deviation.

30. Krausz, F. & Ivanov, M. Attosecond physics. *Rev. Mod. Phys.* **81**, 163–234 (2009).
31. Henke, B., Gullikson, E. & Davis, J. X-ray interactions: photoabsorption, scattering, transmission, and reflection at $E = 50$ –30,000 eV, $Z = 1$ –92. *At. Data Nucl. Data Tabl.* **54**, 181–342 (1993).
32. Domke, M., Schulz, K., Remmers, G., Kaendli, G. & Wintgen, D. High-resolution study of $1P^o$ double-excitation states in helium. *Phys. Rev. A* **53**, 1424–1438 (1996).
33. Santra, R., Yakovlev, V. S., Pfeifer, T. & Loh, Z.-H. Theory of attosecond transient absorption spectroscopy of strong-field-generated ions. *Phys. Rev. A* **83**, 033405 (2011).
34. Samson, J. A. R., He, Z. X., Yin, L. & Haddad, G. N. Precision measurements of the absolute photoionization cross-sections of He. *J. Phys. B* **27**, 887 (1994).
35. Hicks, P. J. & Comer, J. Ejected electron spectroscopy of autoionizing states excited by low energy electron impact. *J. Phys. B* **8**, 1866 (1975).
36. Burgers, A., Wintgen, D. & Rost, J.-M. Highly doubly excited S states of the helium atom. *J. Phys. B* **28**, 3163–3183 (1995).
37. Argenti, L. & Moccia, R. K-matrix method with B-splines: sigma (nl), beta(n) and resonances in He photoionization below $N=4$ threshold. *J. Phys. B* **39**, 2773–2790 (2006).
38. Argenti, L. & Moccia, R. Helium 2(3)S photoionization up to the $N=5$ threshold. *J. Phys. B* **41**, 035002 (2008).

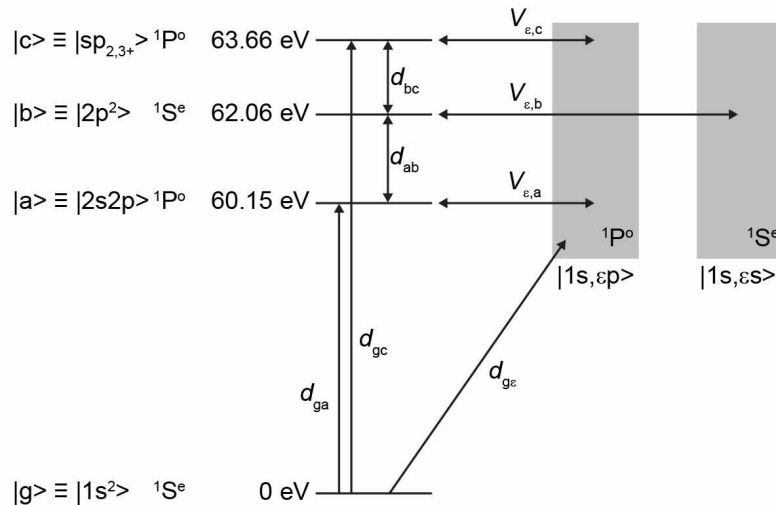


Extended Data Figure 1 | Experimental apparatus and recorded data.

a, Design view of the experimental set-up, consisting of a neon (Ne) gas target for high-harmonic generation (HHG), a motorized iris aperture, a split mirror (SM) in combination with a thin silicon nitride (Si_3N_4) membrane and an aluminium (Al) filter, a focusing toroidal mirror (TM), a dense (~ 100 mbar) absorbing helium (He) target, and a home-built high-resolution spectrometer, which consists of a variable-line-spacing (VLS) grating, a cooled (-50°C) XUV CCD camera, and a pair of Al filters for stray-light suppression. **b**, Recorded

XUV reference spectrum (black line; no He gas in target cell) in the 50–70 eV energy range, averaged over $\sim 64,000$ laser shots, and recorded XUV signal spectrum after transmission through the dense He gas target (red line), averaged over $\sim 640,000$ laser shots. The statistical error is of the order of the line thickness. **c**, Two-dimensional absorbance at a calibrated VIS peak intensity of $3.3 \times 10^{12} \text{ W cm}^{-2}$. The plot consists of 300 single absorbance spectra (for details and definition, see ‘Experimental data acquisition’ in Methods), that were obtained with a time-delay step size of ~ 170 as.

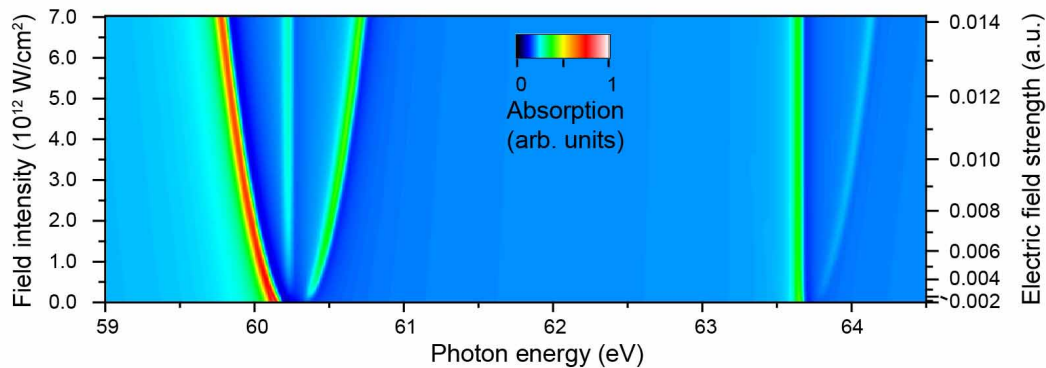
a



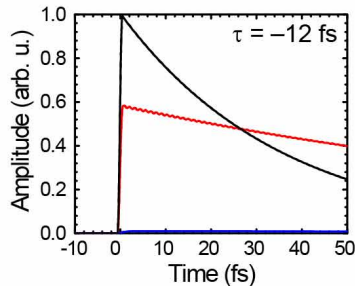
b

$$i\partial_t \begin{pmatrix} c_g \\ c_a \\ c_b \\ c_c \\ c_{\epsilon p} \\ c_{\epsilon s} \end{pmatrix} (t) = \begin{pmatrix} E_g & 0 & 0 & 0 & 0 & 0 \\ d_{ga}^* \cdot F_{XUV}^*(t) & E_a & d_{ab} \cdot F_{VIS}(t) & 0 & V_{\epsilon,a} & 0 \\ 0 & d_{ab}^* \cdot F_{VIS}^*(t) & E_b & d_{bc} \cdot F_{VIS}(t) & 0 & V_{\epsilon,b} \\ d_{gc}^* \cdot F_{XUV}^*(t) & 0 & d_{bc}^* \cdot F_{VIS}^*(t) & E_c & V_{\epsilon,c} & 0 \\ d_{ge}^* \cdot F_{XUV}^*(t) & V_{\epsilon,a} & 0 & V_{\epsilon,c} & [p + A_{VIS}(t)]^2/2 - i\gamma & 0 \\ 0 & 0 & V_{\epsilon,b} & 0 & 0 & [p + A_{VIS}(t)]^2/2 - i\gamma \end{pmatrix} \begin{pmatrix} c_g \\ c_a \\ c_b \\ c_c \\ c_{\epsilon p} \\ c_{\epsilon s} \end{pmatrix} (t)$$

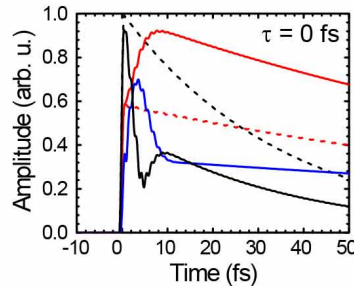
c



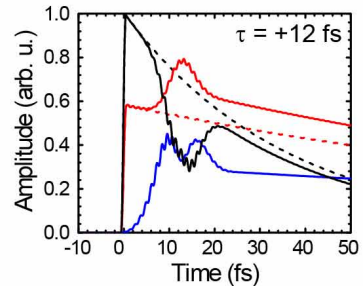
d



e

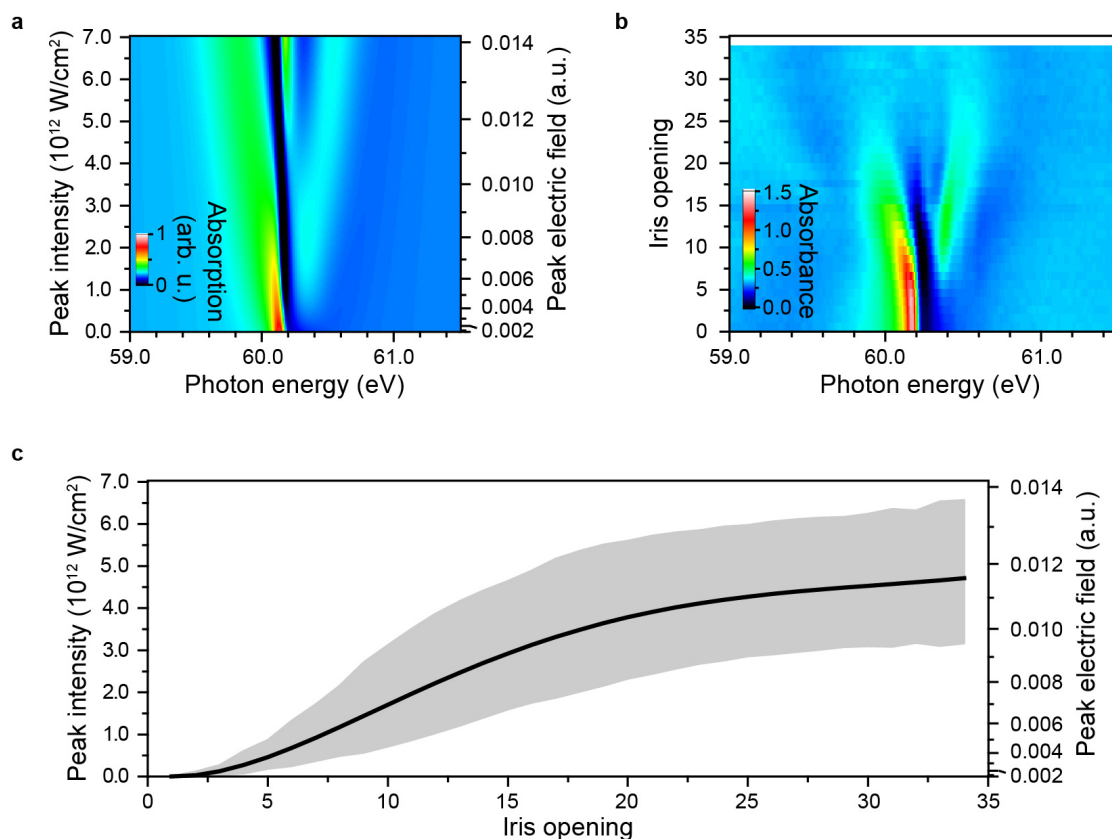


f



Extended Data Figure 2 | Few-level model simulation details. **a**, Level scheme of the simulated subsystem, including the ground state $|g\rangle \equiv |1s^2\rangle$, the autoionizing bound states $|a\rangle \equiv |2s2p\rangle$, $|b\rangle \equiv |2p^2\rangle$ and $|c\rangle \equiv |sp_{2,3+}\rangle$, and the continua $|1s, \epsilon p\rangle$ and $|1s, \epsilon s\rangle$, all coupled via the dipole matrix elements d_{nm} as depicted. The configuration-interaction matrix elements $V_{\epsilon,n}$ couple the excited states with their corresponding (symmetry $^1P^o$ or $^1S^e$) continua. **b**, Schrödinger equation describing the temporal evolution of the coupled states' expansion coefficients $c_n(t)$, resulting from the respective coupling pathways depicted in **a**. Further explanations and definitions of parameters are given in 'Few-level model simulation' in Methods. **c**, Simulated

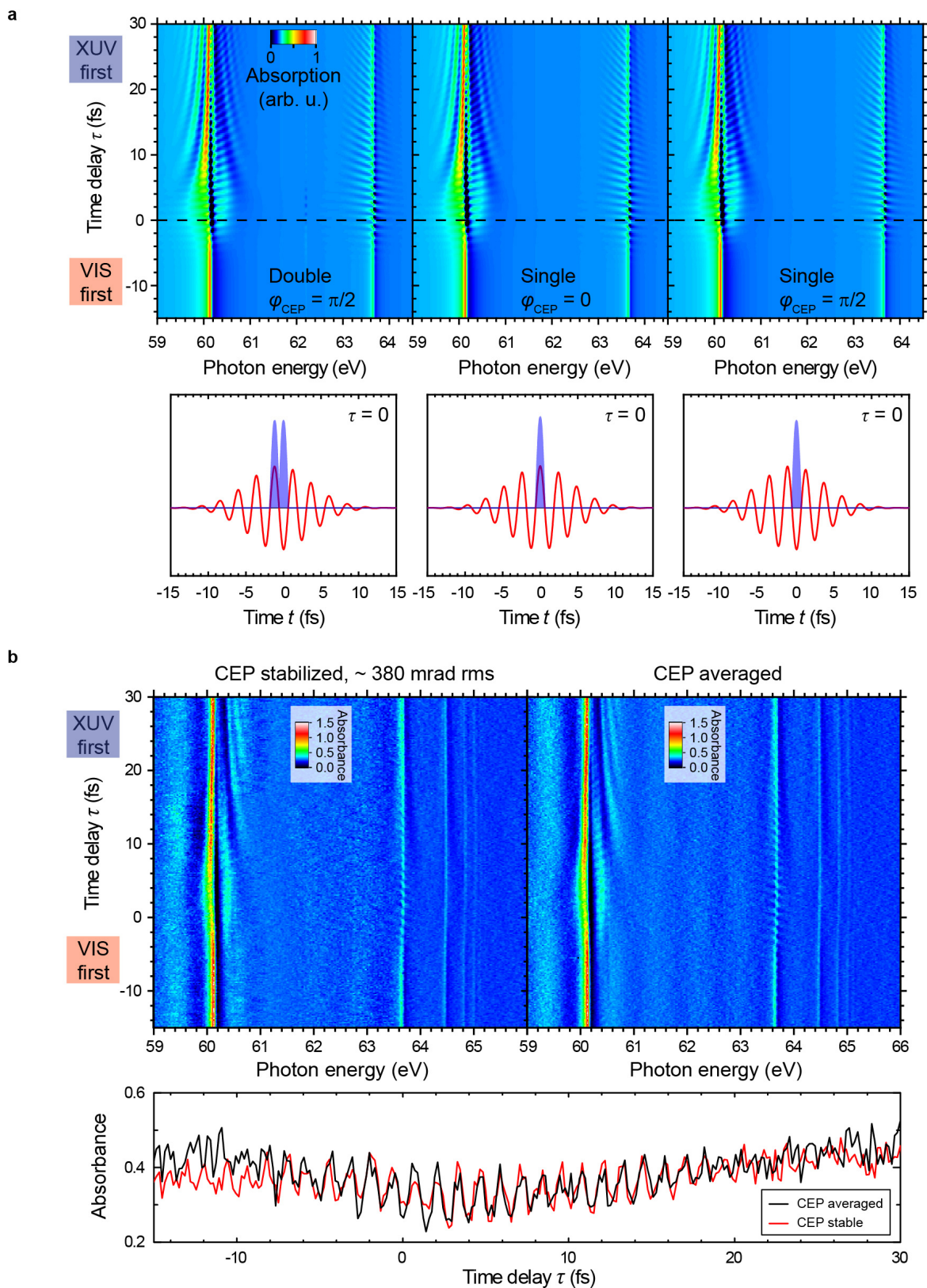
two-dimensional absorbance plot of the few-level system assuming a quasi-monochromatic VIS field of 730 nm wavelength. The absorbance spectra were temporally averaged over one VIS laser cycle (XUV/VIS delay), and convolved with the experimental detector resolution ($\sigma = 20$ meV). **d–f**, Simulated temporal evolution of $|c_n(t)|$ of the three autoionizing states $2s2p$ ($^1P^o$; black lines), $2p^2$ ($^1S^e$; blue lines) and $sp_{2,3+}$ ($^1P^o$; red lines) where the $^1P^o$ -symmetry states were weakly populated by an XUV attosecond pulse at time $t = 0$ fs. All states were coupled by a VIS pulse (7 fs, 730 nm, 3×10^{12} W cm $^{-2}$) at three different time delays τ . The dashed curves show the states' evolution in the absence of the VIS field.



Extended Data Figure 3 | Intensity calibration of the experimental data.

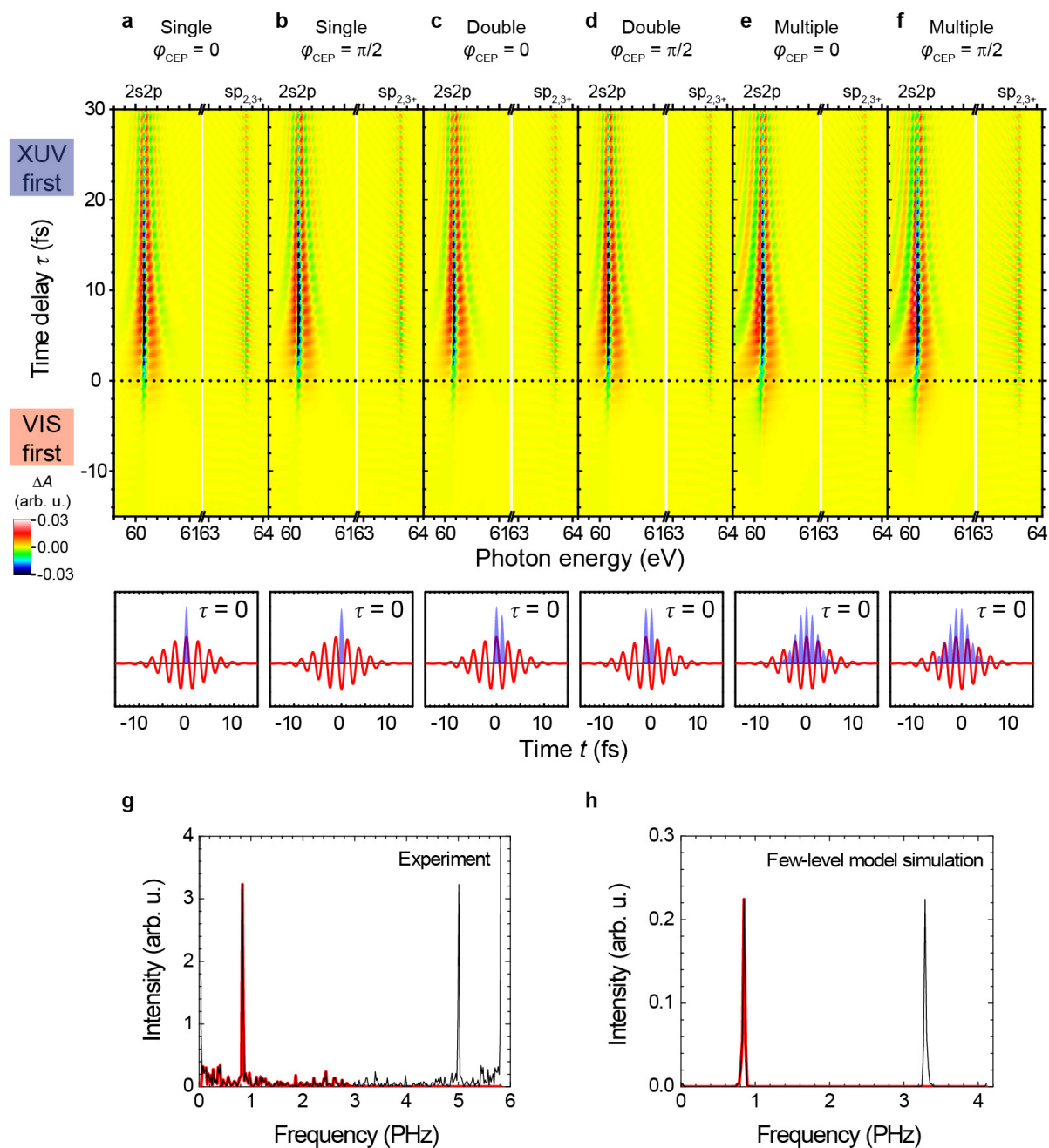
a, Calculated absorbance for a 7 fs, ~730 nm VIS laser pulse at increasing intensity. **b**, Experimentally measured absorbance for increasing openings of the iris diaphragm. For **a** and **b**, the time delay was set to where the Autler-Townes splitting is at maximum, averaged over two modulation periods. **c**, The comparison of maximum absorbance of the left-shifting $2s2p$ line (starting

from 60.15 eV) between numerical and experimental results yields an *in situ* mapping between the VIS intensity and the iris opening in the experiment (black line). The grey shaded area denotes the standard deviation, taking into account different VIS durations and an additional comparison near 0 fs time delay, and thus represents the systematic uncertainty of the monotonically increasing intensity-calibration curve.



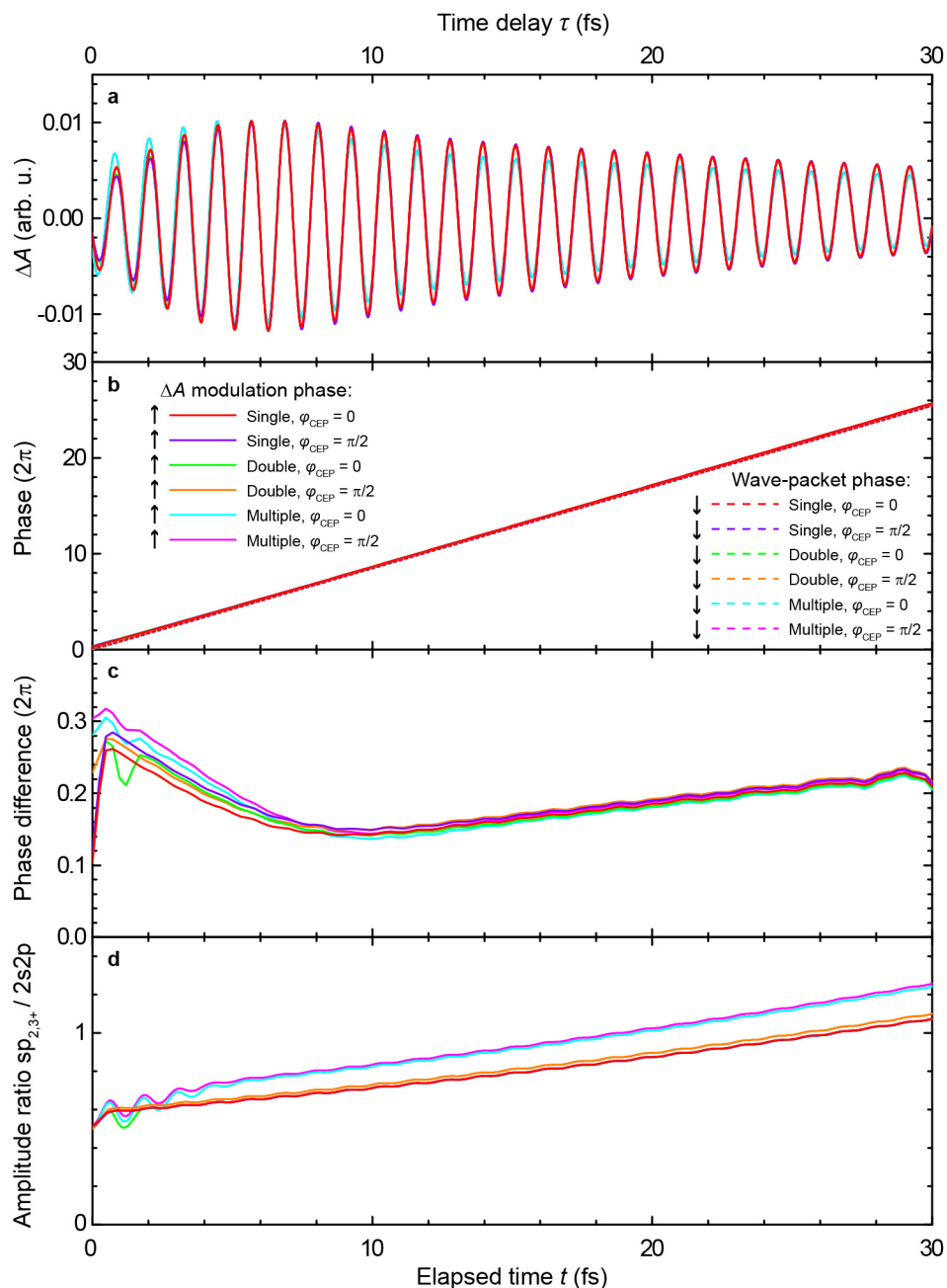
Extended Data Figure 4 | Effects of the attosecond pulse configuration and the carrier envelope phase. **a**, Simulated absorbance plots (top) for different XUV pulse configurations: two attosecond pulses, $\varphi_{\text{CEP}} = \pi/2$ (left); one attosecond pulse, $\varphi_{\text{CEP}} = 0$ (middle); one attosecond pulse, $\varphi_{\text{CEP}} = \pi/2$ (right). The VIS pulse duration was 7 fs with $3 \times 10^{12} \text{ W cm}^{-2}$ peak intensity, where the respective XUV/VIS pulse configurations are illustrated at zero time delay (bottom). **b**, Experimentally measured absorbance plots for CEP stabilization (top left; with root mean squared residual statistical noise of 0.38 rad) and

CEP non-stabilization (top right). The observed time-dependent features, including the lineout at photon energy 63.66 eV (bottom) are practically identical for the CEP-stabilized and the non-CEP-stabilized measurements. Any significant temporal jitter between the attosecond pulses and the VIS carrier wave in the HHG process, for the case of statistical CEP, would correspond approximately to an averaging over a range of time delays for the case of a CEP-stable measurement, smearing out subcycle oscillations in the absorbance. This is clearly not observed.



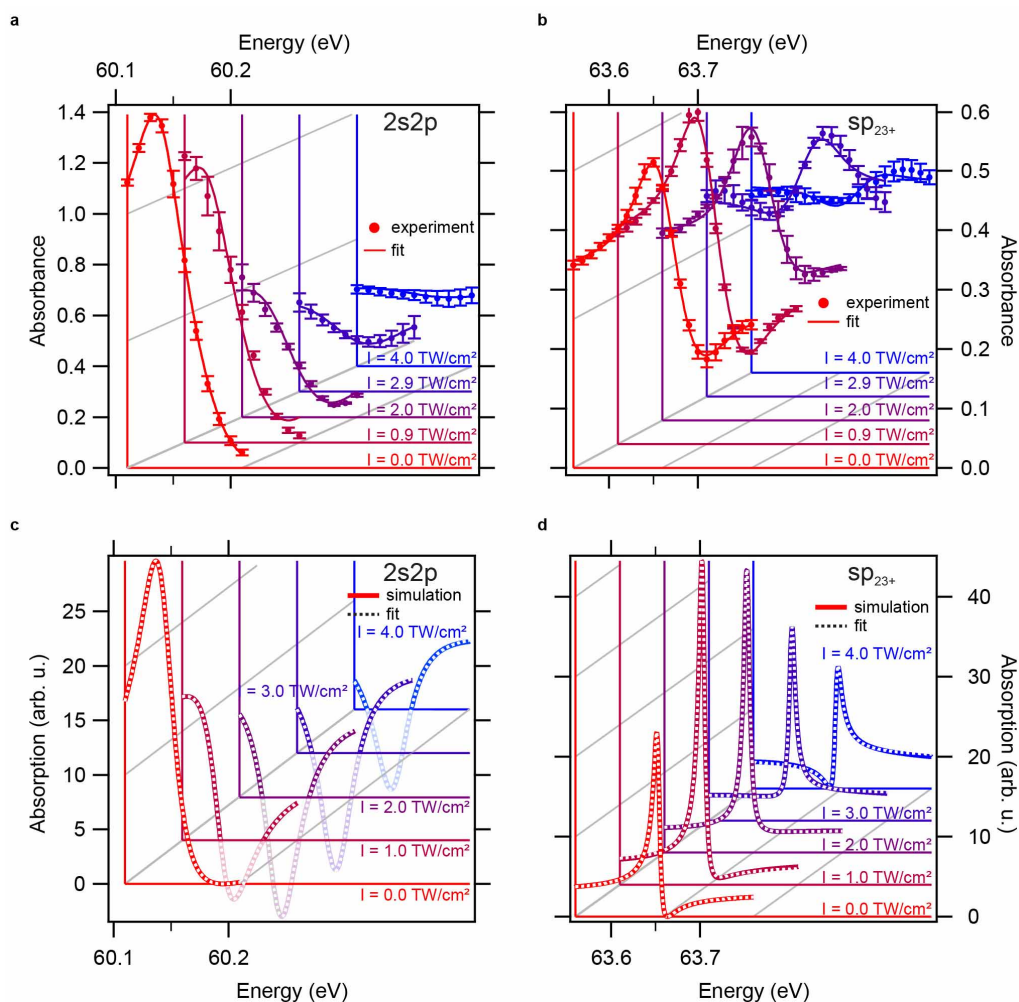
Extended Data Figure 5 | Simulated absorbance changes (ΔA) for low VIS intensity and different pulse configurations. **a**, One attosecond pulse, $\varphi_{\text{CEP}} = 0$. **b**, One attosecond pulse, $\varphi_{\text{CEP}} = \pi/2$. **c**, Two attosecond pulses, $\varphi_{\text{CEP}} = 0$. **d**, Two attosecond pulses, $\varphi_{\text{CEP}} = \pi/2$. **e**, Multiple attosecond pulses (pulse train), $\varphi_{\text{CEP}} = 0$. **f**, Multiple attosecond pulses (pulse train), $\varphi_{\text{CEP}} = \pi/2$. The lower plots in **a–f** show the respective pulse configurations at zero time

delay. The VIS pulse duration was 7 fs and the intensity was $3 \times 10^{10} \text{ W cm}^{-2}$. **g**, **h**, Power spectral density distribution of the ΔA oscillation of the experiment (**g**) and the simulation (**h**). The frequency range used in the analysis is marked in red. We used the full modulation bandwidth, via filtering from near-zero frequency up to the Nyquist frequency, to retrieve the phase $\varphi_A(\tau)$.



Extended Data Figure 6 | Reconstruction of the wave packet from time-delay-dependent $\Delta A(\tau)$ data near the $sp_{2,3+}$ resonance at 63.67 eV (simulation results). **a**, $\Delta A(\tau)$ for the different excitation scenarios shown in Extended Data Fig. 5. **b**, The phase $\varphi_A(\tau)$ extracted from the $\Delta A(\tau)$ oscillations (solid lines), compared with the phase $\varphi(t)$ of the wave packet (dashed lines) for the different excitation configurations. **c**, The difference between the (measurable) modulation phase $\varphi_A(\tau)$ and the wave-packet phase $\varphi(t)$ for each

of the excitation scenarios. A time-delay-dependent correction phase of $\sim 0.4\pi$ needs to be taken into account to reconstruct the wave-packet phase $\varphi(t)$ in the experiment from the measured $\Delta A(\tau)$, as shown in Fig. 2. **d**, The variation in the wave-packet amplitude ratio for the different excitation configurations. Even in the extreme case of multiple attosecond pulses, the amplitude ratio is well defined within a region of $\pm 10\%$.



Extended Data Figure 7 | Fitting the intensity-dependent spectral line shapes of the $2s2p$ and the $sp_{2,3+}$ resonances. **a, b,** Least-squares fit to the experimentally measured line shape shown in Fig. 4a. The laser-controlled line shape is shown for several laser intensities as given in the figure. Error bars here and phase error in Fig. 4c, d correspond to s.d. obtained by analysing at three values of $\tau = 5.35 \text{ fs}$, $\tau = 4.15 \text{ fs}$, $\tau = 6.55 \text{ fs}$. **c, d,** Least-squares fit to the

theoretically predicted line shape obtained from the *ab initio* simulation results shown in Fig. 4b, also plotted for several laser intensities as denoted in the figure. In all cases, the restricted energy region of the least-squares fit ($2s2p$, 60.11–60.21 eV (**a, c**); $sp_{2,3+}$, 63.56–63.76 eV (**b, d**)) ensures phase retrieval for times after the interaction with the laser pulse.

The contribution of the Precambrian continental lithosphere to global H₂ production

Barbara Sherwood Lollar¹, T. C. Onstott², G. Lacrampe-Couloume¹ & C. J. Ballentine³

Microbial ecosystems can be sustained by hydrogen gas (H₂)-producing water–rock interactions in the Earth's subsurface and at deep ocean vents^{1–4}. Current estimates of global H₂ production from the marine lithosphere by water–rock reactions (hydration) are in the range of 10¹¹ moles per year^{5–9}. Recent explorations of saline fracture waters in the Precambrian continental subsurface have identified environments as rich in H₂ as hydrothermal vents and seafloor-spreading centres^{1,2} and have suggested a link between dissolved H₂ and the radiolytic dissociation of water^{10,11}. However, extrapolation of a regional H₂ flux based on the deep gold mines of the Witwatersrand basin in South Africa¹¹ yields a contribution of the Precambrian lithosphere to global H₂ production that was thought to be negligible (0.009 × 10¹¹ moles per year)⁶. Here we present a global compilation of published and new H₂ concentration data obtained from Precambrian rocks and find that the H₂ production potential of the Precambrian continental lithosphere has been underestimated. We suggest that this can be explained by a lack of consideration of additional H₂-producing reactions, such as serpentinization, and the absence of appropriate scaling of H₂ measurements from these environments to account for the fact that Precambrian crust represents over 70 per cent of global continental crust surface area¹². If H₂ production via both radiolysis and hydration reactions is taken into account, our estimate of H₂ production rates from the Precambrian continental lithosphere of 0.36–2.27 × 10¹¹ moles per year is comparable to estimates from marine systems.

Ancient saline fracture waters in the Precambrian continental subsurface, with groundwater residence times ranging from millions¹³ to billions of years¹⁴, provide a previously underestimated source of H₂ for the terrestrial deep biosphere. Until now, little of the information on H₂ in these settings, accessed via underground research laboratories and mines, has been incorporated into global geochemical and biogeochemical models. Figure 1a documents (to our knowledge) the continental sites worldwide for which detailed H₂ studies have been published, as well as new data from our own research sites on the Precambrian Shield in Canada and South Africa (Table 1 and the source data for Fig. 1). Figure 1a shows that the high levels of H₂ reported for the Witwatersrand basin in South Africa by Lin *et al.*¹ are by no means a unique phenomenon. Sites in Precambrian terrains globally have H₂ concentrations as high as those reported for the Witwatersrand basin and for marine hydrothermal systems (Fig. 1a). Notably, sites on the Canadian and Fennoscandian Precambrian Shields and at Phanerozoic ophiolite seeps (such as in Luzon, Semail and Sonoma) and gas wells intersecting kimberlites (in Kansas), have even higher H₂ levels (>30% by volume) than those reported for the Witwatersrand basin (Table 1, Fig. 1a). The notable exposures of ultramafic and mafic rock at many of these sites are consistent with hydration of mafic/ultramafic rocks providing an additional source of H₂ at these sites above and beyond the H₂ produced by radiolysis. Drawing on this global data set, we provide, for the first time, estimates of global H₂ production for the Precambrian continental lithosphere that consider H₂ production from both radiolysis and hydration reactions.

When estimating radiolytic H₂ production, the ratio of H₂ to He (Fig. 1b) can provide important constraints, because He is an inert and conservative tracer. Using measured U, Th and K concentrations, natural α , β and γ particle fluxes can be estimated. Assuming a water-filled porosity of 0.1% and bulk rock density of 2.5 g cm⁻³, Lin *et al.*¹⁰ calculated radiolytic H₂ production rates in water ranging from 10⁻⁸ to 10⁻⁹ nM s⁻¹ for granite, basalt and quartzite lithologies. The radiogenic ⁴He production can also be estimated from U and Th abundances, allowing the H₂/He for radiolytic production of H₂ to be modelled (details in Methods). Since both radiogenic ⁴He and radiolytic H₂ are correlated with U, Th and K concentrations, the H₂/He ratio is, for any given porosity, relatively insensitive to mineralogical composition (felsic, mafic or ultramafic), but is sensitive to porosity changes.

Here we use the Precambrian shield surface area of 1.06 × 10⁸ km² (versus 1.48 × 10⁸ km² for the total continental surface area) to scale the middle and upper continental crust ⁴He production rate¹⁵ to provide an estimate of the ⁴He production rate for the Precambrian crust. Using H₂/He ratios as a function of porosity, we estimate H₂ production from radiolysis for the Precambrian continental lithosphere (details of all calculations in Methods). Estimates of porosity¹⁶ vary from 1.6%–2% in the near surface, down to 0.2% at 10 km and 0.03% at 20 km, averaging 0.96% between 0 km and 10 km (upper crust) and 0.12% between 10 km and 20 km (middle crust). Calculated H₂/He values from radiolysis using the same assumptions as Lin *et al.*¹¹ yield average H₂/He values for the upper and middle continental crust of 117 and 15 for average porosities of 0.96% and 0.12%, respectively. Multiplying the ⁴He production rate by the modelled H₂/He ratios yields a total radiolytic H₂ production rate in the water-filled fractures of the Precambrian crust of 0.16 × 10¹¹ mol yr⁻¹ (Table 2). This is a minimum estimate as it does not include basement rock fluid inclusions that also provide for at least an additional 1% water-filled porosity (see, for example, ref. 17). The latter would produce a further 0.31 × 10¹¹ mol yr⁻¹ to give a total Precambrian crustal H₂ production rate of at least 0.47 × 10¹¹ mol yr⁻¹ (Table 2). This is probably conservative given that if we used typical porosity values published for crystalline rock (up to 2%; details in Methods) rather than 0.12% to 0.96%, this value could be as high as ~1 × 10¹¹ mol yr⁻¹. However, the key point is that even before considering H₂ production via hydration reactions, our estimate of H₂ from the Precambrian continental rocks based on radiolysis alone is similar to marine estimates (Table 2).

Since the discovery of H₂- and CH₄-rich fluids at the Lost City hydrothermal vents in the mid-Atlantic Ocean^{18,19}, there has been increasing interest in the role of water–rock reactions producing energy for chemosynthetic microbial communities, both in marine systems proposed to be analogues of the development of early biosynthetic pathways²⁰, and in continental Phanerozoic ophiolites^{4,21}. Table 2 provides the estimates for high-temperature venting at the mid-ocean ridges^{6,7}, hydration reactions at vents and slow-spreading ridges^{6,8,9,22}, and Fe and sulphide oxidation of basaltic crust⁵—each of which are of the order of 10¹¹ mol yr⁻¹. Although most of these studies refer to their estimates as H₂ flux, it would be more accurate to consider these to be H₂ production rates, as in fact only one of these⁸ is strictly based on a diffusion flux model. The

¹Department of Earth Sciences, 22 Russell Street, University of Toronto, Toronto M5S 3B1, Canada. ²Department of Geosciences, Guyot Hall, Princeton University, Princeton, New Jersey 08544, USA.

³Department of Earth Sciences, South Parks Road, University of Oxford, Oxford OX1 3AN, UK.

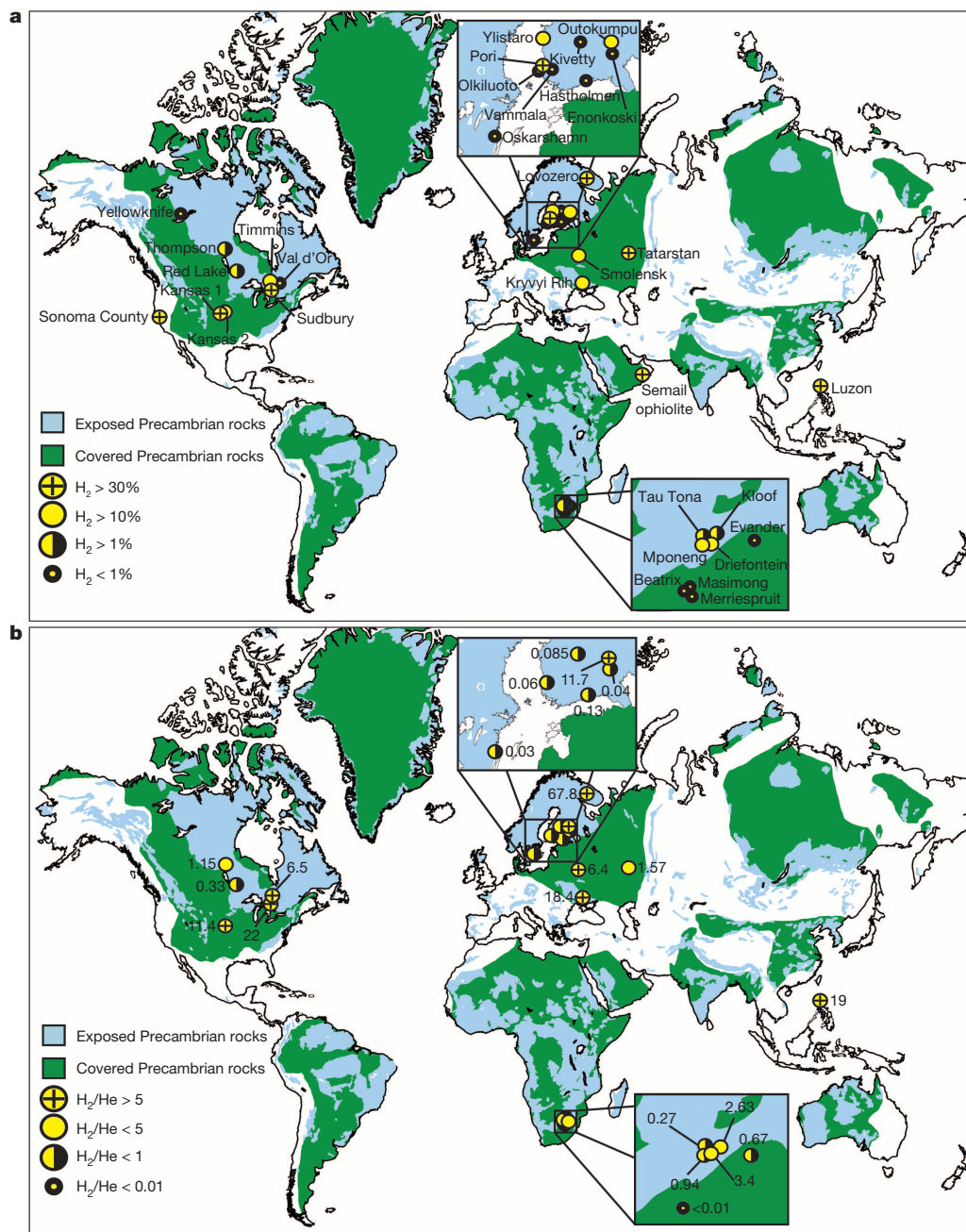
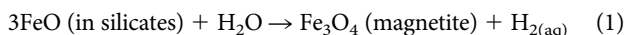


Figure 1 | Precambrian rocks of the continental crust. Geologic data are from ref. 12. Total Precambrian crust, exposed (blue) and buried (green), accounts for $1.06 \times 10^8 \text{ km}^2$, or >70% of total continental crust surface area¹². Symbols for each site show the highest reported H₂ levels in volume per cent (a) and H₂/He ratios (b), with locations provided in Table 1. H₂ concentrations

and H₂/He ratios listed are the maximum observed at each location, but represent a minimum estimate owing to simultaneous H₂ consumption both by microbial communities of sulphate-reducers and methanogens¹ and reaction of H₂ to produce abiogenic hydrocarbons via Fischer–Tropsch synthesis^{21,28}. Map generated via open source software from ref. 29.

others follow the approach typical for this literature, using reaction-based models with a governing equation relating oxidation of FeO in the crust to H₂ production such as the following from Sleep and Bird⁶:



Depending on the study, FeO contents are assumed to be between 5% and 10%, reaction efficiencies range from 100% to the more conservative estimate⁶ of 50%, and typical rock densities fall in the range 3,000–3,500 kg m⁻³. These estimates from the marine lithosphere yield H₂ production estimates in moles per square metre of surface area for oceanic crust to an assumed depth of 1 km. These reaction model calculations are

then coupled to estimates of hydrothermal fluid or water circulation through the spreading centres or ocean floor⁷ or to ocean crust production rates^{5,6} to introduce a temporal term and finally express H₂ production in terms of moles per year (Table 2 and Methods).

To provide the most relevant comparison to the marine literature, we took a similar reaction model approach (detailed calculations in Methods). The iconic greenstone belts of the Precambrian, named for the coloration of the mafic/ultramafic minerals, formed initially as island arcs, continental margin arcs, submarine plateaus, oceanic islands and in some cases, Archean oceanic crust²³. Owing to fundamental changes in the nature of volcanism and heat flux²⁴, the greatest production (and thickness) of greenstone terrains are in the Archean, although formation continued

Table 1 | Maximum H₂ and H₂/He at Precambrian sites and selected Proterozoic ophiolites

Site Name	Location	Latitude	Longitude	H ₂	H ₂ /He
Red Lake	Canada	51° 4' 20.9'' N	93° 46' 45.9'' W	1.60%	0.33
Sudbury	Canada	46° 29' 33.8'' N	81° 0' 38.4'' W	57.8%	22
Thompson	Canada	55° 44' 51.0'' N	97° 51' 4.7'' W	2.75%	1.15
Timmins	Canada	48° 28' 42.2'' N	81° 19' 55.3'' W	12.7%	6.5
Val D'Or	Canada	48° 6' 26.8'' N	77° 47' 11.1'' W	0.51%	NA
Yellowknife	Canada	62° 27' 46.2'' N	114° 22' 41.1'' W	<0.01%	NA
Enonkoski	Finland	62° 5' 22.0'' N	28° 54' 57.9'' E	0.04%	0.04
Hastholmen	Finland	60° 1' 44.0'' N	24° 9' 23.0'' E	1.1%	0.13
Kivetty	Finland	62° 50' 7.8'' N	25° 39' 2.6'' E	0.001%	0.085
Olkiluoto	Finland	61° 14' 19.0'' N	21° 28' 33.0'' E	0.11%	0.06
Outokumpu	Finland	62° 43' 34.1'' N	29° 0' 58.7'' E	12.8%	11.7
Pori	Finland	61° 29' 12.1'' N	21° 47' 53.5'' E	30.4%	NA
Vammala	Finland	61° 20' 28.5'' N	22° 54' 34.8'' E	<0.01%	NA
Ylistaro	Finland	62° 56' 25.8'' N	22° 30' 47.1'' E	11.4%	NA
Beatrix	S. Africa	27° 58' 28.9'' S	26° 44' 4.2'' E	<0.01%	<0.01
Driefontein	S. Africa	26° 23' 60.0'' S	27° 30' 0.0'' E	10.3%	3.4
Evander	S. Africa	26° 24' 59.0'' S	29° 4' 59.5'' E	0.01%	0.67
Kloof	S. Africa	26° 20' 51.3'' S	27° 37' 27.1'' E	9.25%	2.63
Masimong	S. Africa	27° 55' 60.0'' S	26° 45' 0.0'' E	<0.01%	NA
Merriespruit	S. Africa	28° 6' 60.0'' S	26° 51' 0.0'' E	<0.01%	NA
Mponeng	S. Africa	26° 25' 30.0'' S	27° 24' 60.0'' E	11.5%	0.94
TauTona	S. Africa	26° 23' 60.0'' S	27° 24' 60.0'' E	2.40%	0.27
Oskarshamn	Sweden	57° 24' 51'' N	16° 39' 55'' E	0.08%	0.03
Lovozero	Russia	67° 51' 2.5'' N	35° 5' 58.2'' E	35.2%	67.8
Tatarstan	Russia	55° 12' 59.4'' N	50° 45' 11.4'' E	96.1%	1.57
Smolensk	Russia	54° 47' 6.9'' N	32° 3' 1.6'' E	13.5%	6.49
Kryvyi Rih	Ukraine	47° 55' 0.0'' N	33° 15' 0.0'' E	23.2%	18.4
Kansas 1	USA	38° 48' 19.7'' N	96° 52' 5.8'' W	80.0%	NA
Kansas 2	USA	39° 56' 17.4'' N	95° 30' 20.7'' W	17.2%	11.4
Sonoma	USA	39° 5' 45.5'' N	122° 26' 20.3'' W	51.7%	NA
Semail	Oman	20° 36' 44.0'' N	55° 58' 57.6'' E	99.0%	NA
Luzon	Philippines	16° 38' 48.2'' N	121° 15' 54.9'' E	42.6%	19

The table shows the maximum reported H₂ concentrations (in volume per cent of total gas phase) and maximum observed H₂/He ratios (see text) for each site in the Precambrian subsurface shown in Fig. 1, with values for boreholes intersecting kimberlites from Kansas, USA; and samples from younger, surface-exposed Phanerozoic ophiolites (Sonoma County, California; Semail, Oman; and Luzon, Philippines) shown for comparison. In addition to >50 new boreholes/samples published here for the first time, ~150 other boreholes/samples have been compiled from the literature in order to provide a quantitative global context for this phenomenon. For individual data points see the source data for Fig. 1 and for site locations and geologic descriptions see Methods. NA, not analysed.

more rarely throughout the later geologic record²³. Precambrian greenstone sequences can be many kilometres thick^{23,25} and thus differ fundamentally from the Phanerozoic continental ophiolites (relatively thin splinters of oceanic crust obducted onto the continents) that have been the focus of most H₂ production studies so far. Of the total surface area of the continents (1.48×10^8 km²; ref. 12), exposed Precambrian crust, including the uplifted exposed cratons (shown as the blue-shaded areas in Fig. 1), accounts for approximately 30% of the total continental surface area. Including both exposed cratons (blue) and Precambrian crust beneath consolidated Phanerozoic sediments (green-shaded areas in Fig. 1) the total Precambrian crust accounts for 72% of the continents (or 1.06×10^8 km²; ref. 12). Using this value for the total Precambrian continental crust, and based on estimates from ref. 12 that 86% is Proterozoic (9.12×10^7 km²) and 14% is Archean (1.48×10^7 km²), and using the 25% and 50% of mafic/ultramafic rock abundance for Proterozoic and Archean crust respectively²⁶, using a depth of 1 km, we obtain a combined Precambrian rock volume with H₂ production potential via hydration reactions of 3.02×10^{16} m³ (Extended Data Table 1).

Table 2 | Estimates of H₂ production from water-rock reactions

System	H ₂ production (10 ¹¹ mol yr ⁻¹)	Reference
Ocean crust	0.8 to 1.3	Ref. 7
Ocean crust	1.9	Ref. 6
Ocean crust	2.0	Ref. 9
Slow-spreading ridges	1.67	Ref. 8
Basaltic ocean crust	4.5 ± 3.0	Ref. 5
Continental Precambrian radiolysis	0.16 to 0.47	This study
Continental Precambrian hydration reactions	0.2 to 1.8	This study

The table shows global estimates of H₂ production from water-rock alteration reactions (in units of 10¹¹ mol yr⁻¹) from marine lithosphere and H₂ production estimates from radiolysis and hydration of mafic/ultramafic rocks from Precambrian continental lithosphere derived in this study. Estimates made using conservative assumptions. For details of all calculations see Methods. Volcanic, mantle-derived or microbial sources of H₂ are not incorporated.

Although the total thickness of the continental crust is between 30 km and 50 km, we based our estimate on a depth of 5 km, the estimated depth of the habitable zone⁶ (details in Methods). Assuming a rock density of 3,000 kg m⁻³, an average FeO of 10% for these mafic/ultramafic rocks (based on the values of 9.2% to 11.3% given by ref. 27), and a FeO:H₂ ratio of 3:1 as in equation (1), and incorporating the age of the rock, we obtain an estimate of $0.78\text{--}1.8 \times 10^{11}$ mol yr⁻¹ H₂ from the mafic/ultramafic Precambrian crust (Extended Data Table 2).

A more conservative lower boundary can be calculated by incorporating the variation in reaction efficiency. Rather than using 100% reaction efficiency ($\Delta\xi = 1$), as above, we used a second approach that assumes $\Delta\xi = 1/2$ in the uppermost kilometre, $\Delta\xi = 1/4$ in the second kilometre, $\Delta\xi = 1/8$ in the third kilometre, and so on as in ref. 6. Applying this series to the H₂ production rate of $0.78\text{--}1.8 \times 10^{11}$ mol yr⁻¹ produces a minimum estimate of $0.2\text{--}0.4 \times 10^{11}$ mol yr⁻¹ (details in Methods). These upper (1.8×10^{11}) and lower (0.2×10^{11}) boundary estimates are provided in Table 2. Even these are probably conservative, because alternative methods of calculating production rate—based on incorporation of exhumation rates or on extrapolation of published experimental rates of H₂ generation via hydration reactions—all yield estimates of H₂ production that are even higher (details in Methods).

These findings all support the major conclusion of this paper that H₂ production from the Precambrian continental lithosphere, hitherto assumed to be negligible, is in fact an important source of H₂ production. Although H₂ estimates from marine systems provide an important end-member, we suggest that a thorough assessment of the global H₂ potential for supporting a deep subsurface biosphere should not neglect the Precambrian terrain. H₂ production from either radiolysis or hydration of mafic/ultramafic rocks alone revises upward previous published estimates of global H₂ production. The initial estimates provided here suggest that incorporation of H₂ production from the Precambrian continental lithosphere could double existing estimates of global H₂ production from these processes that have been based on marine systems alone.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 July; accepted 27 October 2014.

- Lin, L.-H. *et al.* Long-term sustainability of a high-energy, low-diversity crustal biome. *Science* **314**, 479–482 (2006).
- Sherwood Lollar, B. *et al.* Hydrogeologic controls on episodic H₂ release from Precambrian fractured rocks—energy for deep subsurface life on Earth and Mars. *Astrobiology* **7**, 971–986 (2007).
- D'Hondt, S. *et al.* Subseafloor sedimentary life in the South Pacific gyre. *Proc. Natl Acad. Sci. USA* **106**, 11651–11656 (2009).
- Schrenk, M. O., Brazelton, W. J. & Lang, S. Q. Serpentinization, carbon, and deep life. *Rev. Mineral. Geochem.* **75**, 575–606 (2013).
- Bach, W. & Edwards, K. J. Iron and sulfide oxidation within the basaltic ocean crust: implications for chemolithoautotrophic microbial biomass production. *Geochim. Cosmochim. Acta* **67**, 3871–3887 (2003).
- Sleep, N. H. & Bird, D. K. Niches of the pre-photosynthetic biosphere and geologic preservation of the Earth's earliest ecology. *Geobiology* **5**, 101–117 (2007).
- Canfield, D. E., Rosing, M. T. & Bjerrum, C. Early anaerobic metabolisms. *Phil. Trans. R. Soc. Lond. B* **361**, 1819–1836 (2006).
- Cannat, M., Fontaine, F. & Escartin, J. Serpentinization and associated hydrogen and methane fluxes at slow spreading ridges. *AGU Geophys. Monogr. Ser.* **188**, 241–264 (2010).
- Kasting, J. F. & Canfield, D. E. in *Fundamentals of Geobiology* (eds Knoll, A. H., Canfield, D. E. & Konhauser, K. O.) Ch. 7, 93–104 (Blackwell, 2012).
- Lin, L.-H., Slater, G. F., Sherwood Lollar, B., Lacrampe-Couloume, G. & Onstott, T. C. The yield and isotopic composition of radiolytic H₂, a potential energy source for the deep subsurface biosphere. *Geochim. Cosmochim. Acta* **69**, 893–903 (2005a).
- Lin, L.-H. *et al.* Radiolytic H₂ in the continental crust: nuclear power for deep subsurface microbial communities. *Geochem. Geophys. Geosyst.* **6**, Q07003 (2005b).
- Goodwin, A. M. *Principles of Precambrian Geology* (Academic, 1996).
- Lippmann-Pipke, J. *et al.* Neon identifies two billion year old fluid component in Kaapvaal Craton. *Chem. Geol.* **283**, 287–296 (2011).
- Holland, G. *et al.* Deep fracture fluids isolated in the crust since the Precambrian. *Nature* **497**, 357–360 (2013).
- Ballentine, C. J. & Burnard, P. G. Production, release and transport of noble gases in the continental crust. *Rev. Mineral. Geochem.* **47**, 481–538 (2002).
- Bethke, C. M. A numerical model of compaction-driven groundwater flow and heat transfer and its application to the paleohydrology of intracratonic sedimentary basins. *J. Geophys. Res.* **90**, 6817–6828 (1985).
- Nordstrom, D. K., Lindblom, S., Donahoe, R. J. & Barton, C. C. Fluid inclusions in the Stripa granite and their possible influence on the groundwater chemistry. *Geochim. Cosmochim. Acta* **53**, 1741–1755 (1989).
- Kelley, D. S. *et al.* A serpentinite-hosted ecosystem: the Lost City Hydrothermal Field. *Science* **307**, 1428–1434 (2005).
- Proskurowski, G. *et al.* Abiogenic hydrocarbon production at Lost City Hydrothermal Field. *Science* **319**, 604–607 (2008).
- Lang, S. Q. *et al.* Microbial utilization of abiogenic carbon and hydrogen in a serpentinite-hosted system. *Geochim. Cosmochim. Acta* **92**, 82–99 (2012).
- Etioppe, G. & Sherwood Lollar, B. Abiotic methane on Earth. *Rev. Geophys.* **51**, 276–299 (2013).
- Sleep, N. H., Meibom, A., Fridriksson, T., Coleman, R. G. & Bird, D. K. H₂-rich fluids from serpentinization: geochemical and biotic implications. *Proc. Natl Acad. Sci. USA* **101**, 12818–12823 (2004).
- Condie, K. C. Greenstones through time. In *Archean Crustal Evolution* (ed. Condie, K. C.) Ch. 3, 85–120 (Elsevier, 1994).
- Bickle, M. J. Implications for melting for stabilization of the lithosphere and heat loss in the Archean. *Earth Planet. Sci. Lett.* **80**, 314–324 (1986).
- de Wit, M. & Ashwal, L. D. *Greenstone Belts* (Oxford Monographs on Geology and Geophysics Vol. 35, Clarendon Press, 1997).
- Condie, K. C. Chemical composition and evolution of the upper continental crust: contrasting results from surface samples and shales. *Chem. Geol.* **104**, 1–37 (1993).
- Rudnick, R. L. & Fountain, D. M. Nature and composition of the continental crust: a lower crustal perspective. *Rev. Geophys.* **33**, 267–309 (1995).
- Onstott, T. C. *et al.* Martian CH₄: sources, flux and detection. *Astrobiology* **6**, 377–395 (2006).
- Chorlton, L. B. *Generalized Geology of the World: Bedrock Domains and Major Faults in GIS format: A Small-scale World Geology Map with an Extended Geological Attribute Database* Open File 5529 (Geological Survey of Canada, 2007).

Acknowledgements The preparation and execution of this work was supported by the Canada Research Chairs programme, NSERC Discovery and Accelerator grants to B.S.L. with additional partial funding from the Sloan Foundation Deep Carbon Observatory, Canadian Space Agency and National Science Foundation grant number EAR-0948659.f. We are grateful to K. Chu, A. Yang and G. S. Lollar (of the University of Toronto) for preparation of the maps and tables and to N. Sleep, H. D. Holland, J. Mungall and M. A. Hamilton for discussions on Precambrian geology and mineralogy. We also thank colleagues and supporters at the mines and underground research laboratories whose efforts resulted in the original primary publications from which a portion of this data set is compiled.

Author Contributions B.S.L. designed the project and wrote the paper. B.S.L., C.J.B. and T.C.O. developed the models for H₂ generation. All co-authors contributed to the interpretation and final version of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.S.L. (bslollar@chem.utoronto.ca).

METHODS

In Table 1, for each site, the maximum measured H_2 (percentage of total gas) and H_2/He are reported. Maximum H_2/He ratios provide a conservative (minimum) measure of H_2 production, given that there has probably been loss of H_2 to biological and chemical sinks relative to inert $He^{1,11,21,22,28}$, as has also been noted by studies for the marine H_2 sub-seafloor biosphere where measured H_2 concentrations were low or below detection limit for many samples^{3,30}. Specific measurements of H_2 and H_2/He for each of the >200 samples/boreholes, including 56 previously unpublished, are provided in the source data for Fig. 1. Data are for gases discharging from exploration boreholes in subsurface mines at 19 Precambrian Shield sites from Canada, Finland and South Africa^{2,31–34}. Originally dissolved in saline groundwater in sealed fracture systems in the rocks, gases are released via depressurization into mine workings at rates of 1 to >30 litres of gas per minute per borehole^{14,34}. For comparison, samples from 13 additional sites are included^{13–15} for a total of 32 sites worldwide (see source data for Fig. 1). Although marine and groundwater systems can typically report all measurements as dissolved moles per litre, the database in this paper is drawn from degassing boreholes, in some cases from gas seeps with no corresponding water flow, or from historic data, as well as from fluid inclusions results. The data are all therefore reported as volume per cent of the total gas phase as the only commonly available unit. Sampling methods are described in the relevant publications for each site, and for the mine boreholes in Canada, Finland and South Africa in refs 14 and 33. Compositional analyses of gas samples were performed after the methods of ref. 33. All analysis were run in triplicate and mean values are reported. Reproducibility for triplicate analyses is $\pm 5\%$. Additional details of the geologic settings, sampling methods and analytical methods are provided in the specific references for each site (listed above and in the source data for Fig. 1).

Previous radiolytic H_2 estimates for Precambrian continents. Natural emission of α , β and γ particles released due to decay of U, Th and K was calculated by Lin *et al.*¹⁰ for representative granite, basalt and quartzite lithologies as the basis for calculating radiolytic H_2 production. For each lithology a water-filled porosity of 0.1% and bulk rock density of 2.5 g cm^{-3} was used and stopping powers of 1.5, 1.25 and 1.14 were used for α , β and γ particles respectively. Lin *et al.*¹⁰ calculated the rate of H_2 accumulation within the water from radiolysis (that is, the H_2 production rate in water) to be 9.0×10^{-8} , 9.4×10^{-9} and $2.6 \times 10^{-8} \text{ nM s}^{-1}$, respectively, on the basis of typical ranges of U, Th and K contents for granite, basalt and quartzite lithologies. H_2 production rates of approximately $10^{-8} \text{ nM s}^{-1}$ were reported for a range of felsic lithologies, while for a range of U, Th and K contents typical of mafic and ultramafic lithologies^{26,44} values of $10^{-9} \text{ nM s}^{-1}$ were calculated¹⁰.

Lin *et al.*¹¹ used a steady-state diffusive flux model to calculate a regional flux of H_2 from radiolysis out of the topmost 20 km of the Witwatersrand basin. Using the estimates of 4He production and H_2 production for the different stratigraphic formations of the Witwatersrand basin, as described above, assuming a water-filled effective porosity of 1% and complete interconnection of pore space, they used a steady-state diffusion model, the diffusion coefficient for H_2 in water⁴⁵, and a dC/dz term based on calculating concentration gradients between the stratigraphic formation thickness of the Witwatersrand basin units, to derive a concentration C versus depth z profile for dissolved H_2 in the water. From this they estimated a regional flux specific to the Witwatersrand basin of $\sim 8 \mu\text{mol m}^{-2} \text{ yr}^{-1}$ of H_2 produced by radiolysis^{10,11}. Given the surface area of the Witwatersrand basin of $5.25 \times 10^{10} \text{ m}^2$ ($\sim 350 \text{ km} \times 150 \text{ km}$), this corresponds to a H_2 diffusive flux of $4.2 \times 10^5 \text{ mol yr}^{-1}$, which, if extrapolated to the surface area of the Precambrian continents ($1.06 \times 10^8 \text{ km}^2$) yields a global H_2 flux estimate of $0.009 \times 10^{11} \text{ mol yr}^{-1}$. Based on these estimates, and on the prevailing assumption that radiolysis is the sole H_2 -generating mechanism, the contribution of the Precambrian continental lithosphere to global H_2 production has typically been neglected (for example, in ref. 6), since global estimates from alteration of oceanic crust are typically two orders of magnitude higher (Table 2).

It is unlikely that the Witwatersrand basin can be considered truly steady state in terms of diffusion. The diffusion models on which the Lin *et al.*¹¹ model was based were developed for sedimentary basins rather than crystalline fractured rock. An inherent limitation is using a regional estimate such as this (dependent on the specific formation thicknesses and H_2 concentrations of one regional basin) to extrapolate to a global estimate. Most importantly, by focusing only on H_2 dissolved in fracture waters (H_2 production in waters), this estimate neglects any 4He and H_2 stored in the lithological formations. Lin *et al.*¹¹ assumed that all of the 4He was released from the mineral phases to the pore water, after the method of ref. 46. It is necessarily therefore an underestimate of overall radiolytic H_2 production in these rocks. Here we attempt to address this by focusing instead on developing an estimate based on lithology production rates for H_2 (production rate within a given volume of lithology).

Radiolytic H_2 estimates for Precambrian continents. In this study, we discuss how the ratio of 4He to radiolytic H_2 production changes as a function of depth in the crust, and we use the continental 4He production rate and calculated $^4He/H_2$ ratios to estimate radiolytic H_2 production from the continental crust. The contributions of

H_2 to both the water-filled fracture porosity, and to storage in the form of fluid inclusions, are included in this approach (see main text). The production of 4He and radiolytic production of H_2 in the continental crust are due to the radioelements U and Th (and K, which produces H_2 but does not produce 4He). For a given porosity, H_2 and 4He production rates scale to the radioelement concentrations of the host rock. Hence production rates for H_2 and 4He are correlated, and H_2/He ratio and H_2 lithology production rates increase with increasing porosity.

Reported porosities for granite basement rocks range between 0.9% and 2.3% (ref. 47). Bucher and Stober⁴⁸ report a characteristic porosity for basement rocks of 1.0%, while well tests for effective porosities for the Black Forest basement⁴⁹ and from the Canadian Shield⁵⁰ report a range from 0.1% to 2.1%. To incorporate changes in porosity with depth in the crust, the minimum water available for radiolysis in the continental crust can be estimated by assuming a water-filled fracture porosity T that exponentially declines with depth in kilometres z after the models of ref. 16:

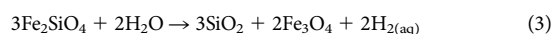
$$T = 1.6e^{-z/4.8} \quad (2)$$

This porosity expression yields a porosity that varies from 1.6% down to 0.2% at 10 km and 0.03% at 20 km, averaging 0.96% between 0 km and 10 km (upper crust) and 0.12% between 10 km and 20 km (middle crust)¹⁶. The porosities predicted by this expression are compatible with He porosities measured in rock units from the Witwatersrand basin⁵¹ and fracture porosities in other crystalline systems^{47–50}. Calculated H_2/He values from radiolysis using the same assumptions as Lin *et al.*¹¹ then yield average H_2/He values for the upper and middle continental crust of 117 and 15 for average porosities of 0.96% and 0.12%, respectively.

The continental crust 4He production rates based on the radioelement content of the upper and middle continental crust are estimated to be $1.8 \times 10^8 \text{ mol yr}^{-1}$ and $1.6 \times 10^8 \text{ mol yr}^{-1}$ (ref. 15). The lower continental crust accounts for $\sim 6\%$ of the 4He (ref. 15) and we neglect this portion of the crust in these calculations. Given that $\sim 70\%$ of the remaining continental crust is comprised of Precambrian basement¹², the production of H_2 from the upper and middle Precambrian shield is $1.2 \times 10^8 \text{ mol yr}^{-1}$ and $1.1 \times 10^8 \text{ mol yr}^{-1}$. Using the H_2/He values calculated above for the fracture porosity (117 and 15 respectively) yields H_2 production in the upper and middle crust of $1.44 \times 10^{10} \text{ mol yr}^{-1}$ and $1.64 \times 10^9 \text{ mol yr}^{-1}$ to give an initial Precambrian crust fracture porosity H_2 production rate of $0.16 \times 10^{11} \text{ mol yr}^{-1}$ (Table 2).

The fracture porosity estimate alone, however, is an underestimate of the H_2O volume exposed to irradiation, because it does not include the fluid inclusion volume, which for basement rocks is typically at least 1% (for example, ref. 17). Radiolytically produced H_2 has been reported in fluid inclusions⁵² and their migration into the fracture water can occur via solid-state diffusion through the host mineral phase, or episodically through metamorphic/tectonic events⁵³ or inclusion decrepitation via fracture propagation¹³. We assumed that the decreasing density of water with increasing temperature and pressure was offset by increasing salinity so that the water density remained $\sim 1 \text{ g cm}^{-3}$. Calculated H_2/He values from radiolysis in water-filled fluid inclusions following Lin *et al.*¹¹ yields minimum H_2/He values for the upper and middle continental crust of 133 and 135 in this case. As above, using the calculated H_2/He values for the water-filled fluid inclusions yields H_2 production in the upper and middle crust of $0.16 \times 10^{11} \text{ mol yr}^{-1}$ and $0.15 \times 10^{11} \text{ mol yr}^{-1}$ to give a total Precambrian crust fluid inclusion H_2 production rate of $0.31 \times 10^{11} \text{ mol yr}^{-1}$. The sum of the fracture porosity ($0.16 \times 10^{11} \text{ mol yr}^{-1}$) and fluid inclusion estimates ($0.31 \times 10^{11} \text{ mol yr}^{-1}$) produces a calculated H_2 Precambrian crustal production rate of $0.47 \times 10^{11} \text{ mol yr}^{-1}$ (Table 2).

This is a conservative estimate, because the porosity-to-depth relationships used reflect average values of porosity. Using the function from ref. 16, which gives a value of 0.96% average porosity for the upper crust and 0.12% for the lower crust, means that in essence we have taken the average porosity of the crust to be $(1.0 + 0.12)/2 = 0.56\%$. For the more representative estimates of porosity for crystalline rock of 1% to 2% described above, the estimated H_2 production from radiolysis could be as high as $1 \times 10^{11} \text{ mol yr}^{-1}$. Importantly, even before an estimate is incorporated for H_2 production via hydration reactions, this estimate of H_2 from the Precambrian continental rocks based on radiolysis alone is similar to marine estimates (Table 2). **H_2 production from hydration reactions in marine systems.** Several studies have produced estimates of global H_2 production from marine systems, including both volcanic/magmatic sources (not considered here) and H_2 production from the abiogenic water–rock alteration reactions that are the focus of this paper. Table 2 provides the estimates for high-temperature venting at the mid-ocean ridges^{6,7}, for warm vents and slow-spreading ridges^{6,8,9,22}, as well as for Fe and sulphide oxidation of basaltic crust⁵. Only one of these marine studies⁸ is based on a diffusion model. All the others follow the approach typical for the marine literature, using reaction-based models with a governing equation relating oxidation of FeO in the crust to H_2 production at a ratio of between 3:1 to 2:1 that is either of the form of equation (1) of ref. 6, or of the following form, from refs 5 and 7:



All the marine studies cited in Table 2 use the following assumptions: FeO content typically between 5% and 10%; reaction efficiency $\Delta\xi$ varying from 100% ($\Delta\xi = 1$) to the more conservative estimate from ref. 6 of 50% ($\Delta\xi = 0.5$); and rock density of $3,000\text{--}3,500\text{ kg m}^{-3}$. These marine literature estimates then calculate H_2 production in moles per square metre of surface area for oceanic crust to a depth of typically 1 km and are then coupled to estimates of hydrothermal fluid or water circulation through the spreading centres or ocean floor⁷ or to ocean crust production rates^{5,6} to introduce a temporal term and then express H_2 production in terms of moles per year (Table 2). Two of the cited studies take a slightly different approach, based on empirical measurements (see discussion below)^{8,9}.

Definition of terms for flux versus production estimates. As the details above indicate, the major references cited in the field for marine global H_2 flux estimates are actually based on reaction-based models rather than diffusion models in many cases. Each of the above papers calculated global H_2 flux from marine water–rock reaction models, as outlined above. In this sense, only ref. 8 can strictly be called a flux estimate, yet—importantly—all the cited papers used the term ‘ H_2 flux’ to describe the results of their reaction-based models. Only ref. 5 use what is likely to be the more appropriate term: ‘production rate’. To provide the best basis of comparison to the marine estimates, we took a similar reaction-based approach for the calculation of H_2 production in this study, as outlined in detail below. For consistency, we have also chosen to use ‘production rate’ to refer to both the previously published marine estimates and the continental Precambrian estimates resulting from this study (Table 2).

Definition of terms for H_2 production via hydration reactions. Although the term serpentinization specifically refers to reaction of the Mg-rich olivine end-member (Fo) to produce serpentine, brucite and magnetite, it is in fact widely used as an umbrella term encompassing a suite of reactions all of which produce H_2 as a by-product of hydration of mafic and ultramafic minerals^{4,8,21,54}. Recent experimental work supports this, demonstrating that H_2 production can occur under a range of temperatures and minerals, including mafic and ultramafic rocks containing peridotite, pyroxene, olivine and magnetite^{55,56}. It is in this sense that the term serpentinization is used here, encompassing the full range of hydration and redox reactions that produce H_2 from alteration of ultramafic and mafic rock.

H_2 production from hydration reactions in Precambrian rocks. In this study, we took a reaction model approach similar to those in the marine literature references described above, using equation (1) from ref. 6, assuming a rock density of $3,000\text{ kg m}^{-3}$, and initially an extent of reaction of $\Delta\xi = 1$. It is important to note that the Precambrian continental crust differs substantially from average continental crust in certain pertinent parameters, in particular, the proportion of mafic and ultramafic rocks and hence FeO content. While the Phanerozoic continental crust is composed of <20% mafic/ultramafic rock, the percentage increases in the Precambrian and is estimated at 25% of Proterozoic crustal rock and between 45% and 51% of the Archean crust²⁶ (Extended Data Table 1). In addition, whereas the total continental crust has an average FeO weight per cent of only 6.6%, the more mafic and ultramafic Archean and Post-Archean rocks range between 9.2% and 11.3% (ref. 27).

Of the total surface area of the continents ($1.48 \times 10^8\text{ km}^2$; ref. 12), exposed Precambrian crust (including the uplifted exposed cratons (for example, the Canadian Shield, Kola Peninsula, the Kaapval Craton) shown as the blue shaded areas in Fig. 1) accounts for approximately 30% of the total continental surface area. Including both exposed cratons (blue) and Precambrian crust beneath consolidated Phanerozoic sediments (green shaded areas in Fig. 1) accounts for 72% of the continental crust (or $1.06 \times 10^8\text{ km}^2$; ref. 12). Using this value for the total Precambrian continental crust, and based on estimates from ref. 12, 86% is Proterozoic in age ($9.12 \times 10^7\text{ km}^2$) and 14% is Archean ($1.48 \times 10^7\text{ km}^2$) (Extended Data Table 1). Knowing from ref. 26 that approximately 25% of the Proterozoic is ultramafic/mafic in composition; and approximately 50% of the Archean is ultramafic/mafic, then for a 1 km depth of crust the volume that has H_2 production potential can be estimated to be $2.28 \times 10^{16}\text{ m}^3$ (Proterozoic) and $0.74 \times 10^{16}\text{ m}^3$ (Archean) respectively.

Assuming a rock density of $3,000\text{ kg m}^{-3}$, an average FeO of 10% for these mafic/ultramafic rocks (based on the values of 9.2% to 11.3% from ref. 27), and a FeO: H_2 ratio of 3:1 as per equation (1), then H_2 production from the ultramafic/mafic Precambrian can be calculated as follows:

$$(3 \times 10^6\text{ g m}^{-3} \times 0.1)/(3 \times 71.845)\text{ g mol}^{-1} = 1.4 \times 10^3\text{ moles H}_2\text{ m}^{-3} \quad (4)$$

Over a depth of 1 km, this corresponds to an estimate of 3.19×10^{19} moles H_2 from the Proterozoic crust and 1.04×10^{19} moles H_2 from the Archean crust. Over an estimated habitable zone of 5 km depth⁶, the above estimates scale to 16.0×10^{19} moles H_2 and 5.2×10^{19} moles H_2 respectively (Extended Data Table 1).

Incorporation of temporal component. Estimates of H_2 production from the marine crust typically convert such reaction-based estimates of H_2 production for a given volume of oceanic crust by incorporating time, either by coupling estimates of total moles H_2 produced to the estimated rate of formation of ocean crust, or by coupling estimates of the rate of hydrothermal fluid circulation and heat flux^{5–7}. Fluid circulation

within Precambrian cratons is more difficult to estimate. However, even in the absence of active tectonism in the Precambrian continents, stress-induced fracturing due to erosion and uplift, and penetration by fracture waters will continue to drive hydration reactions at some finite rate. Typical conceptual models for such systems envisage (1) fracture fluids driving local chemical gradients and renewal of reaction zones at the mineral and fracture interfaces⁵⁵; (2) positive feedback mechanisms wherein reaction-driven cracking further increases permeability and reactive surface areas⁵⁷; and (3) episodic H_2 production due to destabilization of mineral surfaces during progressive water–rock reactions^{2,56}. The episodic nature of these fracture and fluid driven processes means that reaction times will necessarily then be smaller than the total age of the rock. In the absence of detailed information on reaction zones and rates in natural systems, however, by using the age of the rocks as a first approximation, we can derive a conservative estimate of rate. Actual rates could only be larger than these estimates.

Extended Data Table 2 takes this approach and derives global estimates of H_2 production rates in moles per year using the range of ages for the Proterozoic (rounded up to one billion years (Gyr), and to a maximum of 2.5 Gyr) and for the Archean (2.5 to 3.8 Gyr). The resulting estimates of H_2 production rates from Archean and Proterozoic mafic/ultramafic rocks range from $0.14 \times 10^{11}\text{ mol yr}^{-1}$ to $1.6 \times 10^{11}\text{ mol yr}^{-1}$ of H_2 for a total global estimate from the Precambrian lithosphere to a depth of 5 km of $0.78 \times 10^{11}\text{ mol yr}^{-1}$ to $1.8 \times 10^{11}\text{ mol yr}^{-1}$ of H_2 (Extended Data Table 2).

Consideration of the extent of reaction, $\Delta\xi$. Considerations of the likely extent of reaction could provide an even more conservative estimate of H_2 production. The above estimates followed the approach of many of the marine studies, assuming 100% reaction of available FeO (that is, the reaction progress variable $\Delta\xi = 1$; ref. 58). In a second approach, rather than assuming 100% reaction, we assume that $\Delta\xi = 1/2$ in the uppermost kilometre, $\Delta\xi = 1/4$ in the second kilometre, $\Delta\xi = 1/8$ in the third kilometre, and so on, as in ref. 6. Applying this series to the H_2 production rates from 5 km of Precambrian crust derived above ($0.78\text{--}1.8 \times 10^{11}\text{ mol yr}^{-1}$) reduces estimates to $0.2\text{--}0.4 \times 10^{11}\text{ mol yr}^{-1}$. This range of estimated rates 0.2×10^{11} (lower boundary) to $1.8 \times 10^{11}\text{ mol yr}^{-1}$ (upper boundary) are the values listed in Table 2.

Erosion and exhumation rates and experimental constraints. It is helpful to explore additional possible approaches for estimating global H_2 production from the Precambrian in order to constrain the estimates discussed above that form the basis for this paper. Two possibilities are to couple the H_2 production based on the reaction-based models either to erosion rates or to the existing (albeit limited) information on experimental rates of H_2 production via hydration reactions^{56,59}.

As noted, ref. 8's estimate of H_2 production from the marine lithosphere was derived using a different approach from those cited in Table 2 and used 16 H_2 production profiles based on actual H_2 measurements and, for a unit length of ridge axis of a given thickness, calculated a flux by introducing a temporal component using an estimated exhumation rate of 1 cm yr^{-1} (ref. 8). Their reasoning is that the exhumation rate provides the rate of consumption of the crust owing to hydration reactions as exhumation drives propagation of fluid penetration and the reaction front to depth⁶. Applying a similar line of reasoning, the rate of exhumation of the Precambrian lithosphere could be used as an alternative way of incorporating time and deriving rates to compare with those in Extended Data Table 2. Estimates of long-term erosion rates for the Precambrian continents range from $\sim 10\text{ }\mu\text{m yr}^{-1}$ to $2.5\text{ }\mu\text{m yr}^{-1}$ (ref. 60). Taking the total surface area of ultramafic and mafic rock for the Proterozoic and Archean of $3.02 \times 10^{13}\text{ m}^2$ (Extended Data Table 1), even the lower estimate of exhumation rates ($2.5\text{ }\mu\text{m yr}^{-1}$) results in an erosion volume of $7.55 \times 10^7\text{ m}^3\text{ yr}^{-1}$. Given the value of 1.4×10^3 moles $\text{H}_2\text{ m}^{-3}$ (from equation (4)), this annual erosional volume results in an estimate of H_2 production of $1.06 \times 10^{11}\text{ mol yr}^{-1}$. Hence, this alternative approach yields an estimate in the same range as we provide in Table 2—suggesting that indeed the estimated values calculated in the current study are conservative—given that using higher exhumation rates (up to $10\text{ }\mu\text{m yr}^{-1}$) would only increase the contribution of H_2 from the Precambrian crust by this approach, to $>4 \times 10^{11}\text{ mol yr}^{-1}$.

Using experimentally derived rates of H_2 production is challenging, given both the paucity of such experiments so far, and the inherent difficulty of extrapolating laboratory-derived rates to natural systems. H_2 generation rates via hydration reactions will vary with mineralogy, temperature–pressure–oxygen fugacity, and the degree of mineral alteration⁶¹. The presence of komatiite (unaltered ultramafic) textures²⁶ indicates the potential for continuing hydration reactions in Precambrian rocks, although H_2 production rates via hydration reactions in these ancient systems will certainly be slower than in less altered Phanerozoic ophiolites or young ocean floor, owing to the lower temperatures of water–rock reactions in the ancient crust.

Nonetheless, it is useful to explore the implications of H_2 production results via low-temperature water–rock reactions from recent studies. Neubeck *et al.*⁵⁹ published CH_4 production rates during weathering of olivine at $30\text{--}70^\circ\text{C}$ of $(2.7\text{--}7.3) \times 10^{-11}$ moles per metre squared per second (with an associated H_2 production rate, at an H_2/CH_4 ratio of 4:1, of approximately $(10.8\text{--}29.2) \times 10^{-11}$ moles per

- metre squared per second). Assuming a surface area per volume of rock of $300 \text{ cm}^2 \text{ per cm}^3$ (after ref. 16), and extrapolating the rates of ref. 59 to the surface area for Precambrian mafic/ultramafic (Extended Data Table 1; $3.02 \times 10^{13} \text{ m}^2$), the estimated H_2 production over 1 km is of the order of $\sim 1 \times 10^{18} \text{ mol yr}^{-1}$ of H_2 . This is orders of magnitude larger than any of the estimates in the current study (Table 2). This again suggests that the values we derived in this study (Table 2) are conservative, because other published rates of H_2 production from hydration of mafic and ultramafic minerals at $T < 100^\circ\text{C}$ report rates^{56,62} even greater than those of ref. 59.
30. Wankel, S. D. *et al.* Influence of subsurface biosphere on geochemical fluxes from diffuse hydrothermal fluids. *Nature Geosci.* **4**, 461–468 (2011).
 31. Sherwood Lollar, B. *et al.* Evidence for bacterially generated hydrocarbon gas in Canadian Shield and Fennoscandian Shield rocks. *Geochim. Cosmochim. Acta* **57**, 5073–5085 (1993a).
 32. Sherwood Lollar, B. *et al.* Abiogenic methanogenesis in crystalline rocks. *Geochim. Cosmochim. Acta* **57**, 5087–5097 (1993b).
 33. Ward, J. A. *et al.* Microbial hydrocarbon gases in the Witwatersrand Basin, South Africa: implications for the deep biosphere. *Geochim. Cosmochim. Acta* **68**, 3239–3250 (2004).
 34. Sherwood Lollar, B. *et al.* Unravelling abiogenic and biogenic sources of methane in the Earth's deep subsurface. *Chem. Geol.* **226**, 328–339 (2006).
 35. Vovk, I. F. in *Saline Water and Gases in Crystalline Rocks* Special Paper 33 (eds Fritz, P. & Frape, S. K.) 197–210 (Geological Society of Canada, 1987).
 36. Potter, J., Rankin, A. H. & Treloar, P. J. Abiogenic Fischer-Tropsch synthesis of hydrocarbons in alkaline igneous rocks: fluid inclusion, textural and isotopic evidence from the Lovozero complex, N.W. Russia. *Lithos* **75**, 311–330 (2004).
 37. Pedersen, K. *Microbial Processes in Radioactive Waste Disposal*. Report TR-00-04 (Swedish Nuclear Fuel and Waste Management Company (SKB), 2000).
 38. Morrill, P. L. *et al.* Geochemistry and geobiology of a present-day serpentinization site in California: the Cedars. *Geochim. Cosmochim. Acta* **109**, 222–240 (2013).
 39. Fritz, P., Clark, I. D., Fontes, J.-C., Whiticar, M. J. & Faber, E. in *Water-Rock Interaction Vol. 1 Low Temperature Environments* (ed Kharaka, Y. & Maest, A. S.) 793–796 (1992).
 40. Neal, C. & Stanger, G. Hydrogen generation from mantle source rocks in Oman. *Earth Planet. Sci. Lett.* **66**, 315–320 (1983).
 41. Abrajano, T. A. *et al.* Geochemistry of reduced gas related to serpentinization of the Zambales ophiolite, Philippines. *Appl. Geochem.* **5**, 625–630 (1990).
 42. Coveney, R. M., Jr, Goebel, E. D., Zeller, E. J., Dreschhoff, G. A. M. & Angine, E. E. Serpentinization and the origin of hydrogen gas in Kansas. *Am. Assoc. Petrol. Geol. Bull.* **71**, 39–48 (1987).
 43. Newell, K. D. *et al.* H_2 -rich and hydrocarbon gas recovered in a deep Precambrian well in Northeastern Kansas. *Nat. Resour. Res.* **16**, 277–292 (2007).
 44. Salters, V. J. M. & Stracke, A. Composition of the depleted mantle. *Geochim. Geophys. Geosyst.* **5**, 1–27 (2004).
 45. Jaehne, B., Heinz, G. & Dietrich, W. Measurement of the diffusion coefficients of sparingly soluble gases in water. *J. Geophys. Res.* **92**, 10767–10776 (1987).
 46. Lippmann, J. *et al.* Dating ultra-deep mine waters with noble gases and ^{36}Cl , Witwatersrand Basin, South Africa. *Geochim. Cosmochim. Acta* **67**, 4597–4619 (2003).
 47. Aquilina, L., de Dreuz, J. R., Bour, O. & Davy, P. Porosity and fluid velocities in the upper continental crust (2 to 4 km) inferred from injection tests at the Soultz-sous-Forets geothermal site. *Geochim. Cosmochim. Acta* **68**, 2405–2415 (2004).
 48. Bucher, K. & Stober, I. Fluids in the upper continental crust. *Geofluids* **10**, 241–253 (2010).
 49. Stober, I. Permeabilities and chemical properties of water in crystalline rocks of the Black Forest, Germany. *Aquat. Geochem.* **3**, 43–60 (1997).
 50. Stober, I. & Bucher, K. Hydraulic properties of the crystalline basement. *Hydrogeol. J.* **15**, 213–224 (2007).
 51. Silver, B. J. *et al.* The origin of NO_3^- and N_2 in deep subsurface fracture water of South Africa. *Chem. Geol.* **294–295**, 51–62 (2012).
 52. Savary, V. & Pagel, M. The effects of water radiolysis on local redox conditions in the Oklo, Gabon natural fission reactors 10 and 16. *Geochim. Cosmochim. Acta* **61**, 4479–4494 (1997).
 53. Lowenstern, J. B., Evans, W. C., Bergfeld, D. & Hunt, A. G. Prodigious degassing of a billion years of accumulated radiogenic helium at Yellowstone. *Nature* **506**, 355–358 (2014).
 54. Charlou, J. L., Donval, J. P., Fouquet, Y., Jean-Baptiste, P. & Holm, N. Geochemistry of high H_2 and CH_4 vent fluids issuing from ultramafic rocks at the Rainbow hydrothermal field ($36^\circ 14' \text{N}$, MAR). *Chem. Geol.* **191**, 345–359 (2002).
 55. Andreani, M., Daniel, I. & Pollet-Villard, M. Aluminum speeds up the hydrothermal alteration of olivine. *Am. Mineral.* **98**, 1738–1744 (2013).
 56. Mayhew, L. E., Ellison, E. T., McCollom, T. M., Trainor, T. P. & Templeton, A. S. Hydrogen generation from low-temperature water-rock reactions. *Nature Geosci.* **6**, 478–484 (2013).
 57. Kelemen, P. B. & Hirth, G. Reaction-driven cracking during retrograde metamorphism: olivine hydration and carbonation. *Earth Planet. Sci. Lett.* **345–348**, 81–89 (2012).
 58. Helgeson, H. C. in *Geochemistry of Hydrothermal Ore* (ed. Barnes, H. L.) 568–610 (Wiley, 1979).
 59. Neubeck, A., Duc, N. T., Bastviken, D., Crill, P. & Holm, N. G. Formation of H_2 and CH_4 by weathering of olivine at temperatures between 30 and 70°C . *Geochim. Trans.* **12**, <http://dx.doi.org/10.1186/1467-4866-12-6> (2011).
 60. Flowers, R. M., Bowring, S. A. & Reiners, P. W. Low long-term erosion rates and extreme continental stability documented by ancient (U-Th)/He dates. *Geology* **34**, 925–928 (2006).
 61. McCollom, T. M. & Bach, W. Thermodynamic constraints on hydrogen generation during serpentinization of ultramafic rocks. *Geochim. Cosmochim. Acta* **73**, 856–875 (2009).
 62. Stevens, T. O. & McKinley, J. P. Abiotic controls on H_2 production from basalt-water reactions and implications for aquifer biogeochemistry. *Environ. Sci. Technol.* **34**, 826–831 (2000).

Extended Data Table 1 | Volumes of mafic/ultramafic rock with H₂ production potential and H₂ production to depths of 1 km and 5 km

Continental Lithosphere	% †	Surface Area (10 ⁷ km ²)	% Mafic & Ultramafic ‡	Volume Mafic & Ultramafic to 1 km (x 10 ¹⁶ m ³)	Moles H ₂ to 1 km (x 10 ¹⁹) §	Moles H ₂ to 5 km (x 10 ¹⁹) §
Proterozoic	86	9.12	25	2.28	3.19	16.0
Archean	14	1.48	50	0.74	1.04	5.2
TOTAL Precambrian *	100	10.60	-	3.02	4.23	21.2

(See Methods for detailed calculations and discussion.)

* Based on total surface area for Precambrian continental lithosphere of 1.06×10^8 km² (ref. 12).

† Based on Proterozoic and Archean surface areas, accounting for 86% and 14%, respectively, of the total Precambrian continental surface area (refs 12 and 27).

‡ Mafic and ultramafic from ref. 26.

§ Based on 1.4×10^5 moles H₂ per cubic metre of mafic and ultramafic rock from equation (4) (Methods) and based on $\Delta\xi = 1$ (100% reaction efficiency).

Extended Data Table 2 | Estimated H₂ production rates from Precambrian mafic/ultramafic rock for a 5 km volume

Continental Lithosphere	Moles H ₂ to 5 km (x 10 ¹⁹) *	Minimum Age (x 10 ⁹ yr)	Moles H ₂ per year (x 10 ¹¹)	Maximum Age (x 10 ⁹ yr)	Moles H ₂ per year (x 10 ¹¹)
Proterozoic †	16.0	1.0	1.6	2.5	0.64
Archean	5.2	2.5	0.2	3.8	0.14
TOTAL Precambrian	21.2	-	1.8	-	0.78

(See Methods for detailed calculations and discussion.)

*From Extended Data Table 1.

† Minimum age for Proterozoic rounded up to 1.0 Gyr. Using values of <1 Gyr would only increase the estimates of H₂ production rates in this table. Throughout this study, we attempted to provide lower boundaries on H₂ production (conservative estimates). Actual production is thus likely to be higher.

Global protected area expansion is compromised by projected land-use and parochialism

Federico Montesino Pouzols^{1†*}, Tuuli Toivonen^{1,2*}, Enrico Di Minin^{1,3}, Aija S. Kukkala¹, Peter Kullberg¹, Johanna Kuusterä^{1,4}, Joona Lehtomäki¹, Henrikki Tenkanen², Peter H. Verburg⁵ & Atte Moilanen¹

Protected areas are one of the main tools for halting the continuing global biodiversity crisis^{1–4} caused by habitat loss, fragmentation and other anthropogenic pressures^{5–8}. According to the Aichi Biodiversity Target 11 adopted by the Convention on Biological Diversity, the protected area network should be expanded to at least 17% of the terrestrial world by 2020 (<http://www.cbd.int/sp/targets>). To maximize conservation outcomes, it is crucial to identify the best expansion areas. Here we show that there is a very high potential to increase protection of ecoregions and vertebrate species by expanding the protected area network, but also identify considerable risk of ineffective outcomes due to land-use change and uncoordinated actions between countries. We use distribution data for 24,757 terrestrial vertebrates assessed under the International Union for the Conservation of Nature (IUCN) ‘red list of threatened species’⁹, and terrestrial ecoregions¹⁰ (827), modified by land-use models for the present and 2040, and introduce techniques for global and balanced spatial conservation prioritization. First, we show that with a coordinated global protected area network expansion to 17% of terrestrial land, average protection of species ranges and ecoregions could triple. Second, if projected land-use change by 2040 (ref. 11) takes place, it becomes infeasible to reach the currently possible protection levels, and over 1,000 threatened species would lose more than 50% of their present effective ranges worldwide. Third, we demonstrate a major efficiency gap between national and global conservation priorities. Strong evidence is shown that further biodiversity loss is unavoidable unless international action is quickly taken to balance land-use and biodiversity conservation. The approach used here can serve as a framework for repeatable and quantitative assessment of efficiency, gaps and expansion of the global protected area network globally, regionally and nationally, considering current and projected land-use pressures.

Habitat loss and fragmentation due to intensifying land-use is one of the major drivers of biodiversity loss^{7,8}. The global protected area (PA) network is one of the most important means to halt such loss^{1–4}. Adoption of the strategic Aichi Biodiversity Target 11 of the Convention of Biological Diversity (CBD; <http://www.cbd.int/sp/targets>) provides a unique opportunity for expanding the current PA network to cover 17% of the terrestrial areas by 2020. At present, global patterns in biodiversity and global priority areas for conservation at the regional scale are relatively well known^{1,6,8,12–16}, but spatial assessments are essential^{13,14} to maximize global conservation outcomes from PA expansion.

Here, we carried out a comprehensive assessment of priority areas for expanding the current global PA network, and quantified their potential contribution to global conservation. We present a prioritization of the global PA network expansion to 17% that shows the performance and spatial pattern of alternative expansions of the current PA network, delivering balanced, complementary coverage across a breadth of ecoregions (827) and species (24,757), for present and future (2040) land-use conditions, and comparing the outcomes of a globally coordinated expansion

against nationally prioritized expansion areas. Our analyses and maps are informative at the global, regional and national levels.

We used newly developed prioritization methods and software that follow principles and approaches from systematic conservation planning and spatial conservation prioritization^{17,18}. As urbanization, agricultural land-use, desertification and deforestation are rapidly increasing^{8,11}, we integrated information about projected land-use change¹¹ and discounted the distributions of species and ecoregions, to produce effective ranges at present and by 2040 (Supplementary Information). We address three questions that are crucial for the effective implementation of the Aichi Biodiversity Target 11: (1) what is the potential performance of the expanded PA network in terms of increased coverage of species ranges and ecoregions; (2) how will land-use change by 2040 effect the performance and spatial pattern of the best PA expansion areas; and (3) what is the efficiency gap between globally and nationally identified priority areas.

First, our results show that there is a high potential to increase coverage of ecoregions and species, which could be harnessed with complementarity-based prioritization. If placed efficiently (Fig. 1 and Extended Data Fig. 1), additional protection could triple the average protection of vertebrate species ranges (Fig. 2, labels A and B, and Extended Data Fig. 2). Furthermore, it would increase average protection of ecoregions by a factor of 3.3, helping to address the continuing biome crisis¹⁹ and providing a broader bioclimatic coverage and representativeness under climate change²⁰ (Supplementary Information). This high potential is a result of the presently largely unprotected status of a considerable proportion of species and ecoregions that have narrow ranges. Globally, the highest priorities for expanding PAs are located in the Neotropics (Central America, along the Andes and the Brazilian coast), Africa (Madagascar, the Eastern Arc Mountains and the forests of west Africa) and southeast Asia (the Himalayan slopes, Indonesia, Papua New Guinea and the Philippines) (Fig. 1 and Supplementary Information). The locations of the top 17% priorities are relatively consistent at a regional scale (Supplementary Information), regardless of the land-use scenario and/or parameters being used. This highlights the importance of the top priority areas and the robustness of our results in identifying some well-known areas^{1,12,13}.

Second, regarding the effects of projected land-use change on the performance of the expanded PA network, we show that intensification may lead to considerable biodiversity loss by 2040 (Fig. 2, label C). Although the expansion to 17% could on average account for ~61% of the current ranges of species and ecoregions (Fig. 2), the level of protection would drop to ~54% by 2040, even if projected land-use change is accounted for in the PA network expansion. Globally, terrestrial vertebrates could lose on average ~12 to 16% of their current effective range by 2040 (Supplementary Information), with more than 50% habitat loss for more than 2,600 species (Extended Data Fig. 3 and Extended Data Table 1). Furthermore, a loss of 15% in the average range of threatened species would occur by 2040 (Extended Data Table 2), and among threatened species, over 1,000 would lose more than 30% of their current range, 440 more than 50%, and 110

¹Finnish Centre of Excellence in Metapopulation Biology, Department of Biosciences, Biocenter 3, University of Helsinki, PO Box 65 (Viikinkaari 1), FI-00014 Helsinki, Finland. ²Department of Geosciences and Geography, University of Helsinki, PO Box 64 (Gustaf Hållströmin katu 2a), FI-00014 Helsinki, Finland. ³Department of Life Sciences, University of KwaZulu-Natal, University Road, Private Bag X54001, Durban 4000, South Africa. ⁴Regional Council of Helsinki-Uusimaa, Esterinportti 2 B, FI-00240 Helsinki, Finland. ⁵Institute for Environmental Studies, VU University Amsterdam, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands. [†]Present address: Rutherford Appleton Laboratory, Science & Technology Facilities Council, Harwell Oxford Campus, Didcot OX11 0QX, UK.

*These authors contributed equally to this work.

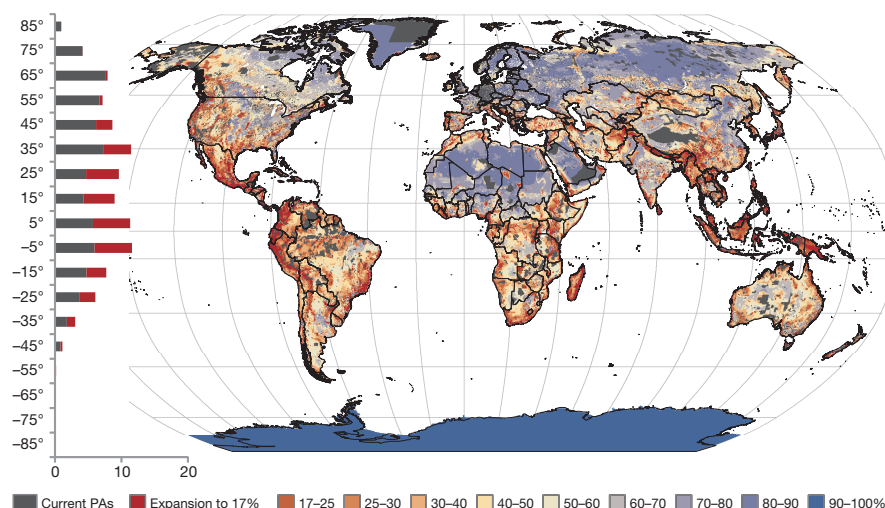


Figure 1 | Global priority map for the expansion of the PA system.

Prioritization of the global PA network expansion, taking future (2040) projected land-use into account. The bars on the left show the distribution of current (grey) and proposed (red) expansion areas by latitude bins. Currently designated PAs are quite evenly distributed across latitudes (55% of global PAs are in latitudes $\geq -30^\circ$ and $\leq 30^\circ$), whereas the expansion effort would be

concentrated in the tropics to maximize coverage of species and ecoregions (75% of the expansion areas are between latitudes -30° and $+30^\circ$). Analysis data sources: International Union for the Conservation of Nature (IUCN), World Database on Protected Areas (WDPA), and Database of Global Administrative Areas (GADM).

more than 70% (Extended Data Table 1 and Supplementary Information). Prioritizing on the basis of threatened species would improve the average coverage of threatened species ranges by only 4%, but causing an average loss of 5% across all non-threatened species and 22% across ecoregions (Extended Data Table 2, Extended Data Fig. 3 and Supplementary Information). Consequently, actions should be taken quickly to reduce land-use changes or palliate their effect in the highest priority areas. Furthermore, to reach the currently possible protection levels, if conservation planning would consider projected future land-use (Fig. 1), the global protection target would need to be increased from 17% to 21% to compensate for land-use intensification (Fig. 2, label C).

Third, we show that global to continental scale conservation planning and international cooperation is vital for reaching high conservation outcomes. We demonstrate this by conducting analyses separately for each country and analysing the resulting global pattern and performance. We find that a lack of international coordination would cause an efficiency loss much greater than expected from projected land-use change by 2040 (Fig. 2 and Extended Data Table 2). The national top 17% areas could at best cover on average $\sim 70\%$ of the amount of species' ranges and ecoregions covered in the global prioritizations (Fig. 2, label C). Although marked overlaps between global and national priorities occur in large tropical countries such as Brazil, Congo and India (Fig. 3 and Extended Data Fig. 4), many highly irreplaceable biodiversity areas in Central America, Madagascar and southeast Asia would be left unprotected in national prioritizations, and over 450 threatened species would lose more than 50% of their effective range (Supplementary Information). Nevertheless, the fraction (38%) in which the global and national priorities overlap (Fig. 3) undoubtedly identifies key areas for Aichi Biodiversity Target 11. In other regions, conservation partnerships across country borders are crucial²¹. This is particularly relevant for the connectivity or compactness of PAs: the global prioritization produces a network in which the number and size of new PAs are comparable to the current network, whereas national prioritization would lead to a more fragmented network, duplicating the number of PAs and decreasing their average size by 60% (see Supplementary Information).

We have made use of several sources of information, including spatial patterns of PAs of all sizes²², high-resolution human-driven land-use scenarios^{11,23}, and spatial patterns of thousands of narrow-range species and distinctive ecoregions^{9,10}. To meet our study objectives, it has been crucial to be able to account for detailed spatial patterns of PAs and

biodiversity (Supplementary Information). Considering the dynamic nature of PA designations, and the numerous downgrading, downsizing and degazettement events recently observed²⁴, there is a need for recurrent following-up on previous studies that have provided insight into the effectiveness and gaps of the global PA network^{1,4,13,15}. Further development of global data resources are required to consider other aspects of biodiversity and additional taxa^{25,26}, such as invertebrates¹⁴ or plants¹⁶. Fine-scale conservation planning assessments using more local information should be carried out in priority areas identified by this study². In particular, high-resolution data can be used on sites of confirmed importance for biodiversity, such as 'important bird areas', 'important plant areas', 'alliance for zero extinction sites', or key biodiversity areas (KBAs) generally² (see Supplementary Information for an analysis of KBAs in three countries). Furthermore, fair estimation of opportunity costs, dynamic monitoring of

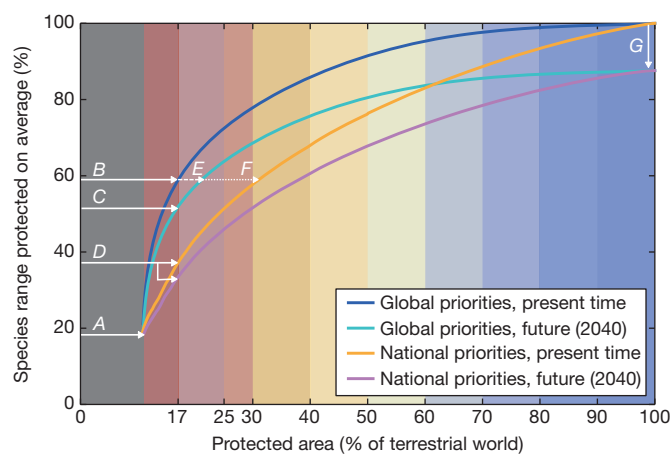


Figure 2 | Cumulative average coverage of species ranges in different fractions of terrestrial land. Terrestrial land fractions are listed in priority order, from current PAs (grey) to 17% expansion (red), and over entire terrestrial land. Background colours match the priority map (Fig. 1). The present PAs cover $\sim 19\%$ of species ranges (A). Expansion to 17% could increase coverage to $\sim 61\%$ (B) or $\sim 56\%$ with 2040 land use (C). National priorities perform much worse (D). A further expansion would be required to compensate land-use change (to 21%, E) and/or national-scale planning (to 32%, F). Globally, land-use change may cause over $\sim 12\%$ species' range loss (G).

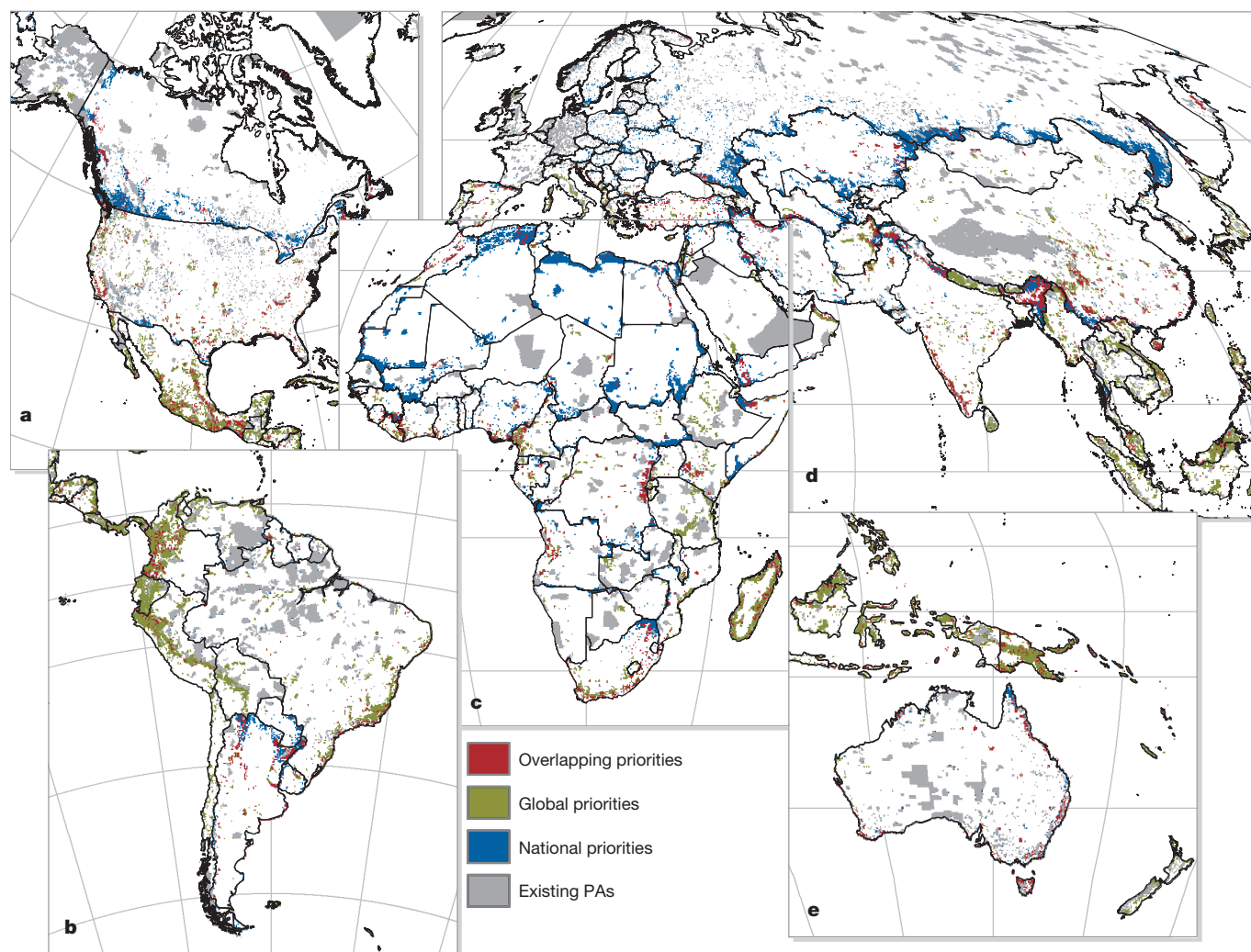


Figure 3 | Global and national priority expansion areas (2040), and their overlap (38% of top 17% priority areas). There is a clear difference between the areas with relatively higher national priority (blue) and higher global priority (green). The edge effects in the national prioritization originate primarily from latitudinal gradients in species diversity. Notably, the congruent areas (red) overlap with many previously identified biodiversity hotspots in

large countries: Atlantic forest/Brazil, Himalaya and mountains of southwestern China and eastern Afrotropical/Congo. While this map is visualized for a strict top 17% threshold, our results provide continuous rankings of the whole land surface of the Earth. For a global map projection, see Supplementary Information. Analysis data sources: IUCN, WDPA and GADM.

threats, update of land-use scenarios, and integration of species-specific habitat requirements would benefit recurrent systematic assessments of biodiversity patterns. While our study focuses on relatively short-term changes due to land-use, climate change is a major issue that needs to be addressed in forward-looking conservation prioritization. For longer-term projections, priorities should be defined considering recent advances in climate change scenarios²⁷, and recent results that model the vulnerability of species²⁸ and ecoregions²⁰ to climate change.

Implementing PA network expansion could be more challenging in areas that are less economically developed, resource limited and/or have weaker governance²⁹. Our global solution (Fig. 1) shows that most of the priority areas for expanding the PA network are concentrated in the global south (Extended Data Fig. 6 and Supplementary Information), whereas only 25% of the global expansion responsibility lies at higher latitudes ($\leq 30^\circ$ and $\geq -30^\circ$). Continentally, Asia has the highest responsibility, with 37% of the total expansion areas, while 18% are in Africa and 31% in Central and South America. In these areas of highest responsibility, support mechanisms are needed to address governance challenges, overall feasibility, development and population growth, and the burden of additional management costs of PAs²⁹. It would also be important to reconcile future land-use with national and global conservation priorities. If every country is to contribute the same percentage of area, priority areas are

more evenly distributed globally, less concentrated in Central and South America and more in Africa and Asia, less in tropical forests and more in temperate forest, and especially in grassland, savannah and shrubland (Extended Data Fig. 6 and Supplementary Information).

Robust, reproducible assessments are pivotal for well-informed and iterative decision-making towards an effective and balanced expansion of the global PA network. Our analysis is based on published data, and a well-documented, newly developed, and publicly available dedicated software tool. We have also shared the files required to implement the analyses, and the resulting spatial data layers in the hope that they will stimulate further analyses and interpretation. Here, we have quantitatively shown the considerable potential that is at stake. Halting biodiversity loss requires global planning and implementation of support mechanisms for the PA network expansion. Furthermore, good coverage of species' ranges in PAs does not guarantee their persistence. The effectiveness of PAs depends on several ecological and societal factors. While the national level implementation is not efficient in terms of global coverage of biodiversity, it is socially more acceptable and increases the local benefits of conservation, that is, the several positive aspects of parochialism³⁰. The Aichi Biodiversity Target 11 opens a unique window of opportunity with political commitment to address biodiversity loss. It is important that decision-makers and other stakeholders take action to implement platforms for effective and balanced

protected area expansion at global, continental and regional scales, and use these to reduce land-use pressures on biodiversity.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 June; accepted 7 November 2014.

Published online 14 November 2014.

- Rodrigues, A. S. L. *et al.* Effectiveness of the global protected area network in representing species diversity. *Nature* **428**, 640–643 (2004).
- Butchart, S. H. M. *et al.* Protecting important sites for biodiversity contributes to meeting global conservation targets. *PLoS ONE* **7**, e32529 (2012).
- Thomas, C. D. *et al.* Protected areas facilitate species' range expansions. *Proc. Natl Acad. Sci. USA* **109**, 14063–14068 (2012).
- Le Saout, S. *et al.* Protected areas and effective biodiversity conservation. *Science* **342**, 803–805 (2013).
- Butchart, S. H. M. *et al.* Global biodiversity: indicators of recent declines. *Science* **328**, 1164–1168 (2010).
- Hoffmann, M. *et al.* The impact of conservation on the status of the world's vertebrates. *Science* **330**, 1503–1509 (2010).
- Gibson, L. *et al.* Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature* **478**, 378–381 (2011).
- Laurance, W. F. *et al.* Averting biodiversity collapse in tropical forest protected areas. *Nature* **489**, 290–294 (2012).
- The International Union for the Conservation of Nature. *The IUCN Red List of Threatened Species*. Version 2013. 2 <http://www.iucnredlist.org> (2013).
- Olson, D. M. *et al.* Terrestrial ecoregions of the world: a new map of life on Earth. *Bioscience* **51**, 933–938 (2001).
- van Asselen, S. & Verburg, P. H. Land cover change or land-use intensification: simulating land system change with a global-scale land change model. *Glob. Change Biol.* **19**, 3648–3667 (2013).
- Gaston, K. J. Global patterns in biodiversity. *Nature* **405**, 220–227 (2000).
- Brooks, T. M. *et al.* Global biodiversity conservation priorities. *Science* **313**, 58–61 (2006).
- Kremen, C. *et al.* Aligning conservation priorities across taxa in Madagascar with high-resolution planning tools. *Science* **320**, 222–226 (2008).
- Jenkins, C. N. & Joppa, L. Expansion of the global terrestrial protected area system. *Biol. Conserv.* **142**, 2166–2174 (2009).
- Joppa, L. N., Visconti, P., Jenkins, C. N. & Pimm, S. L. Achieving the convention on biological diversity's goals for plant conservation. *Science* **341**, 1100–1103 (2013).
- Margules, C. R. & Pressey, R. L. Systematic conservation planning. *Nature* **405**, 243–253 (2000).
- Moilanen, A., Wilson, K. A. & Possingham, H. P. *Spatial Conservation Prioritization: Quantitative Methods and Computational Tools* (Oxford Univ. Press, 2009).
- Hoekstra, J. M., Boucher, T. M., Ricketts, T. H. & Roberts, C. Confronting a biome crisis: global disparities of habitat loss and protection. *Ecol. Lett.* **8**, 23–29 (2005).
- Watson, J. E. M., Iwamura, T. & Butt, N. Mapping vulnerability and conservation adaptation strategies under climate change. *Nature Clim. Change* **3**, 989–994 (2013).
- Moilanen, A., Anderson, B. J., Arponen, A., Pouzols, F. M. & Thomas, C. D. Edge artefacts and lost performance in national versus continental conservation priority areas. *Divers. Distrib.* **19**, 171–183 (2013).
- IUCN & UNEP-WCMC. *The World Database on Protected Areas (WDPA)*. <http://www.protectedplanet.net/> (2013).
- Pereira, H. M. *et al.* Scenarios for global biodiversity in the 21st century. *Science* **330**, 1496–1501 (2010).
- Mascia, M. B. & Pailler, S. Protected area downgrading, downsizing, and degazettement (PADDD) and its conservation implications. *Conserv. Lett.* **4**, 9–20 (2011).
- Beger, M. *et al.* Conservation planning for connectivity across marine, freshwater, and terrestrial realms. *Biol. Conserv.* **143**, 565–575 (2010).
- Mokany, K., Harwood, T. D., Overton, J. M., Barker, G. M. & Ferrier, S. Combining α - and β -diversity models to fill gaps in our knowledge of biodiversity. *Ecol. Lett.* **14**, 1043–1051 (2011).
- Moss, R. H. *et al.* The next generation of scenarios for climate change research and assessment. *Nature* **463**, 747–756 (2010).
- Foden, W. B. *et al.* Identifying the world's most climate change vulnerable species: a systematic trait-based assessment of all birds, amphibians and corals. *PLoS ONE* **8**, e65427 (2013).
- McCarthy, D. P. *et al.* Financial costs of meeting global biodiversity conservation targets: current spending and unmet needs. *Science* **338**, 946–949 (2012).
- Hunter, M. L. & Hutchinson, A. The virtues and shortcomings of parochialism: conserving species that are locally rare, but globally common. *Conserv. Biol.* **8**, 1163–1165 (1994).

Supplementary Information is available in the online version of the paper.

Acknowledgements F.M.P., T.T., E.D.M., A.S.K., P.K., J.L. and A.M. thank the European Research Council Starting Grant (ERC-StG) 260393 (Global Environmental Decision Analysis, GEDA), the Academy of Finland centre of excellence programme 2012–2017 and the Natural Heritage Services (Metsähallitus) for support. P.H.V. thanks the ERC grant 311819 (GLOLAND). We thank A. Santangeli, I. Hanski and H. Tuomisto for comments on the manuscript, and CSC-IT Center for Science Ltd, administered by the Ministry of Education, Science and Culture of Finland, for its support and high-performance computing services. We are grateful for the efforts of data providers, IUCN, BirdLife International, Conservation International, the IUCN Species Survival Commission Specialist Groups and IUCN Red List Partners, the World Wildlife Fund, the United Nations Environment Programme (UNEP) World Conservation Monitoring Centre and the IUCN World Commission on Protected Areas, and their partners and contributors for kindly providing publicly available data, without which this and many other studies would not have been possible.

Author Contributions F.M.P., T.T. and A.M. wrote the manuscript, with contributions from all authors. F.M.P., T.T., E.D.M., J.L., P.K. and A.M. designed the study. A.M. conceived and led the study. F.M.P. and T.T. analysed the data and prepared the figures and tables. F.M.P. implemented prioritization algorithms and analyses. T.T., E.D.M., A.S.K., P.K., J.K., J.L., H.T. and F.M.P. collected and processed the data. P.H.V. contributed land-use models and data. T.T., E.D.M., A.S.K., P.K., J.K., J.L., H.T. and F.M.P. collected and processed the data.

Author Information Original results data and additional interactive visualizations are available online at <http://avaa.tdata.fi/web/cbig/>. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.M.P. (federico.montesino-pouzols@stfc.ac.uk), T.T. (tuuli.toivonen@helsinki.fi) or A.M. (atte.moilanen@helsinki.fi).

METHODS

Below, we describe the steps of our analysis from spatial data collection and pre-processing to spatial prioritization.

Data processing. We based our analysis on a set of 25,588 spatial data layers collected from different sources and in different formats. This data set consisted of two basic administrative delineations (protected areas and country borders), 25,584 biodiversity feature distributions (species and ecoregions) and two land cover layers (present and projected future), as described below. The pre-processing was the same for all layers: we converted all input data to latitude/longitude coordinate system and rasterized them to global resolution grids (latitude/longitude coordinate system), using ArcGIS 10.1 software, harmonized to different resolutions: 0.01667° (equalling 1.7 km at the Equator), 0.05°, 0.1° and 0.2°. The content of the data was restricted to terrestrial areas using a binary land/water mask that contained all continental terrestrial areas, excluding large water bodies³¹. The land/water mask was originally downloaded from and processed by WorldGrids (<http://www.worldgrids.org>).

Basic administrative delineations. The data on protected areas was based on the June 2013 release of the WDPa²² (retrieved from <http://www.protectedplanet.net>, produced by the United Nations Environment Programme's World Conservation Monitoring Centre). We extracted the protected areas from the WDPa database by selecting only areas belonging to IUCN protected area categories I to VI and having as status 'designated' or alike (such as 'designated'). These areas cover approximately 11% of the Earth's land surface (including Antarctica) at the time of this study. We included only protected areas having detailed geographic information in the database (105,369), excluding the ones represented with a point only. This meant excluding in total 21,248 protected areas that did not have polygon boundaries, totalling 817,321 km² (6.9% of all protected areas). One common approach would have been to represent the PAs with only point information by a circle that has the surface of the PA as presented in the WDPa. This would have, however, added extra noise to the shapes of the PAs, as many PAs are elongated or otherwise of particular shape. We rasterized the protected areas to the analysis resolution with an intersect rule, thus labelling all cells touching a protected area polygon as protected areas. This way, we were also able to include the smallest and narrowest protected areas to the analysis. National boundaries were rasterized from the polygons of the GADM based on the unique country codes. This resulted in a raster layer identifying 253 countries or autonomous regions in the world.

Ecoregions. We used spatial distributions of all 827 terrestrial ecoregions, grouped into 14 biomes or major habitat types, as defined by the World Wildlife Fund (<http://worldwildlife.org/biomes>). On the basis of regional analyses and information from hundreds of experts, the ecoregion boundaries delimit areas within which ecological and evolutionary processes interact most strongly¹⁰. The same ecoregion classification has previously been used in analysis of, for example, broad patterns of biodiversity, habitat loss and conservation status of different areas^{15,19,20}.

Species. We based our analysis on terrestrial vertebrates included in the IUCN red list of threatened species^{9,32–35}. Produced by the IUCN Global Species Programme, the IUCN Species Survival Commission, and the IUCN Red List Partnership, this is the most comprehensive global assessment of the conservation status of animal, plant and fungi species. We retrieved the species range data for mammals, amphibians and reptiles from the 'spatial data download' area of the IUCN red list website (<http://www.iucnredlist.org/>). Data for birds was obtained from the BirdLife International data zone web-page³² (<http://www.birdlife.org/datazone/home>).

Distribution data for species were available as geographic information system (GIS) polygons, covering known or inferred areas where species occur. These distribution polygons are in practice positioned somewhere between the extent of occurrence and the true area of occupancy of the species^{36,37}. They are far from perfect and may overestimate the species' true area of occupancy^{4,37–40}, as they may include areas from which the species is absent, such as large freshwater bodies within terrestrial species' distributions. Therefore, the present analyses should be interpreted in terms of coverage of species' ranges, not in terms of coverage of the true distributions of species. Nevertheless, the range maps reduce geographical biases and fill gaps that exist in point locality data^{37,40}. In addition, these species distribution polygons represent the best frequently updated and publicly available information of the distribution limits of vertebrate species⁴. These data have been widely used previously^{4,6,34,35,41–44}. Here, we refined these range distribution maps to obtain effective ranges by land-use models, as further described below. At the time of this study, a considerable fraction of reptiles remains unassessed and range distribution data are not available⁴⁵. The main results reported here have been generated including the available spatial data on reptiles, which are geographically biased⁴⁵. See Supplementary Information for an analysis of the sensitivity of our results to this factor, and prioritization results generated excluding reptiles from the analysis. From the IUCN species database⁹, we selected terrestrial species only, leaving out 79 entirely marine mammals in families Otariidae, Phocidae, Odobenidae, Balaenidae, Balaenopteridae, Delphinidae, Eschrichtiidae, Iniidae, Monodontidae, Neobalaenidae, Phocoenidae, Physeteridae, Platanistidae, Ziphiidae and Sirenia. We processed all species similarly and rasterized the range of each species to a separate raster

layer. With the information facilitated by the IUCN red list of threatened species, the breeding and non-breeding portions of the ranges of migratory birds could also be treated separately⁴⁶.

In the rasterizing process, we assigned the pixel values according to the certainty of species presence in the polygon, as reported by the IUCN. We used four categories with a continuous scale from 1 to 0, with less reliable occurrence categories translated into lower values: extant = 1.0; probably extant and uncertainly extant = 0.5; possibly extinct = 0.1; and extinct = 0.0. Several arguments coming from the field of biogeography suggest the use of spatial resolutions comparable to the highest resolution of the available distribution maps^{37,47,48}. We made the polygon to raster conversion originally using a pixel size of 0.00833° (equalling roughly 0.85 km at the equator) and aggregated the data up to 1.7-km resolution (0.01667°), 0.05°, 0.1° and 0.2° by summing up the original pixel values in blocks of 4, 36, 144 and 576 cells, respectively. This way, we were able to include even the smallest ranges without exaggerating their size.

Land-use data. We considered land-use effects on ecoregion extents and species ranges by discounting the ranges by land use for present time and 2040 (refs 7, 11). This process reduces one of the most common sources of commission errors in species' range maps: areas that fall inside range polygons but are unsuitable for species, as they have been transformed by human activities. For present land-use conditions, species' ranges were discounted by an average of 14.12% (s.d.: 13.37%, median: 9.558%), whereas for future conditions their ranges were discounted by 23.99% on average (s.d.: 19.03%, median: 18.22%). The land-use scenarios for 2040 is based on the Organisation for Economic Co-operation and Development (OECD) environmental outlook baseline scenario⁴⁹. The scenarios were generated using the CLUMondo model at a resolution of 5 arcmin (9.25 km). In the models, land-use changes are driven by regional demand for goods and influenced by local factors that either promote or constrain land-use change. CLUMondo has the highest thematically relevant land-use information for the purpose, distinguishing different land systems that can have a mixed composition and contains relevant information from the perspective of biodiversity analyses. In particular, these models include quantitative information of land-use intensity for different land-use classes¹¹. We first converted the original land-use maps from 2000 (present) and 2040 (future) to numerical data by giving different land-use classes values between 1 and 0 reflecting their naturalness and different intensities of farming^{50,51}. The following naturalness values were given for different land uses, from most to least natural.

Dense forest, mosaic grassland and forest, mosaic grassland/bare and natural grassland = 1.0; open forest/few livestock, open forest, grassland/few livestock, grassland, bare/few livestock = 0.9; mosaic cropland and grassland, mosaic cropland and forest = 0.8; mosaic cropland (extended) and grassland/few livestock, mosaic cropland (extended) and open forest/few livestock = 0.7; mosaic cropland (medium intensive) and grassland/few livestock, mosaic cropland (medium intensive) and forest/few livestock = 0.6; mosaic cropland (intensive) and grassland/few livestock, mosaic cropland (intensive) and forest/few livestock = 0.5; cropland extensive/few livestock, cropland extensive = 0.4; cropland medium intensive/ few livestock, cropland medium intensive = 0.3; cropland intensive/few livestock, cropland intensive = 0.2; bare, peri-urban and villages = 0.1; urban = 0.0. In a more restrictive scale, we defined the naturalness value as 0 for all intensive land uses (see Supplementary Information for additional results obtained for this scale).

To produce estimates of effective ranges for present and future, we multiplied the values in the original species range and ecoregion maps using the naturalness map for present and future, respectively. Technically, the calculations were implemented in zonation by using the condition transformation^{52,53}. In the later analyses for the present and future, we used the respectively transformed sets of distribution layers.

These values were defined as a reasonable first approximation. However, we made several assumptions, and especially we assume the same effects across all taxonomic groups. Refinement of this processing step in a reliable manner would require models and evidence on the effects of different land uses on species, which are only recently becoming available for some taxonomic groups⁵⁴. Alternative approaches used in the literature include the use of habitat suitability models⁵⁵ or habitat classification schemes from the IUCN red list of threatened species⁵⁶ to constrain species' range distributions, or the use of additional data³⁷. Such approaches would probably reduce commission errors resulting from broad range maps but would not be trivial to combine with high-resolution land-use data, and could potentially introduce omission errors and other artefacts resulting from the fact that land-use classes do not match habitat classes.

Spatial prioritization method and process. Priority maps were generated as rankings of landscape elements (cells), iteratively ranked from lowest to highest priority for conservation (Fig. 1). Together with ranking maps, we produced performance curves that describe the extent to which each feature or species is retained in any given high- or low-priority fraction of the landscape (Fig. 2). We implemented priority ranking with the zonation methods and software for spatial conservation planning^{53,58,59}, which produce ranking maps and performance curves as main outputs. We used the newly developed zonation 4 software tool, introducing methods capable of processing problems four or more orders of magnitude bigger than previously possible^{14,21,59,60}, of the

order of 10^4 species or features and 10^9 landscape elements⁵³. Zonation produces a balanced ranking in which balanced denotes that for any given rank level, such as top 17% areas, these areas are complementary and jointly achieve a well-balanced level of representation across all biodiversity features. Complementarity is a key concept in spatial conservation prioritization and it can be loosely defined as a property of the solution that sites work together efficiently in achieving conservation objectives^{4,17,61,62}.

We used the additive benefit function analysis variant of zonation^{53,58,63}, which can be interpreted as minimization of aggregate extinction rates via feature-specific species-area curves. This method can produce a high return on investment⁶⁴ in terms of average coverage of biodiversity features per amount of area protected, and does not require targets or thresholds that necessarily have a degree of arbitrariness⁴¹. To prioritize expansion areas starting from the current global PA network, we used a technique^{14,53,59} in which the priority ranking is generated in two stages, with the ranking of expansion areas being generated in the first stage and current PA landscape elements remaining in the second stage. The method also induces aggregation of cells into compact PAs or PA expansions by favouring cells that are found in the neighbourhood of areas retained for protection to the detriment of more scattered cells⁵³.

The results reported here correspond to seven different set-ups or analysis variants (columns of Extended Data Table 2), which have been made publicly available, with raster maps available for a resolution of 0.2° . Each set-up defines a set of spatial data layers and prioritization analysis parameters based on which a unique priority ranking (together with performance curves) is produced in a deterministic manner. Ecoregions were weighted so that their aggregate weight is equal to the aggregate weight of all species. Species were weighted according to their category of extinction risk on the IUCN red list of threatened species⁶⁵, with highest weights assigned to critically endangered species (least concern: 1, near threatened: 2, vulnerable 4, endangered: 6, critically endangered: 8, data deficient: 2). This weighting scheme induces a relatively higher coverage of more endangered species while the prioritization method maintains an overall balanced representation of different species and groups of species (Supplementary Information). In the seven different prioritization set-ups we analysed the implications of: (1) different land-use conditions (present and future, 2040); (2) whether all assessed species or only threatened species are considered as priorities for conservation; and (3) the context of planning, that is, defining global priorities in a globally coordinated manner versus strictly nationally developed priorities. Alternative analysis variants excluding reptiles were also evaluated. In these, although the data on reptile species' distributions is strongly geographically biased, the figures of global expansion responsibility by latitude change only slightly, with a 0.4% decrease of responsibility in latitudes between 30° and -30° (see Supplementary Information).

The analysis presented here implicitly assumes that costs (acquisition, management and opportunity) of protected areas are uniform across the world, whereas in practice costs vary enormously⁶⁶. Different approaches to integrate costs into conservation planning have been proposed in the literature⁶⁷. Costs can be integrated in a zonation prioritization analysis in different ways⁵³, and global data on conservation costs are publicly available⁶⁸, although these have several limitations⁴⁴. The integration of costs also requires careful consideration of other factors that can have a major influence on spatial conservation prioritization, such as governance⁶⁹, funding issues⁷⁰ or the dynamic nature of other societal factors in a changing world with areas experiencing an increase in public demand for conservation and willingness to pay for conservation, especially in tropical countries⁷¹.

The main results presented here correspond to analyses carried out for input grid layers with a resolution of 0.2° , or approximately 20 km at the equator. This low resolution was used in our main results to reflect the limitations in the original input data on species' distributions, reducing potential misuse of our results. In particular, the data limitations should be carefully considered when making decisions at a local scale. See Supplementary Information for additional analysis results corresponding to different, higher analysis resolutions up to 1.7 km. We found that our results, when aggregated globally, continentally or nationally, or by species groups or latitude bins are robust with respect to the analysis resolution used in the range explored here (from 0.01667° to 0.2° degrees).

The spatial prioritization approach used here uses two kinds of data: distribution data of biodiversity features and costs (where relevant), and structural data elements. The first class includes input data digitized to polygons at various scales. With high resolution it is possible to mimic the shapes of the original species distributions without introducing an additional bias in early analysis stages. The second class of data includes mask layers, such as those defining spatial units, such as country borders and protected area boundaries. These are typically known and digitized as spatial data with high precision.

Accounting for land-use change. Three set-ups were used to analyse the implications of projected land-use change on global priorities for expanding the PA network: global priorities present time, global priorities (2040), and global priorities (restrictive 2040) (Extended Data Table 2). Here and in general, the set-up for present time uses effective ranges of species and extents of ecoregions according to present land-use conditions, whereas the set-up for 2040 uses effective distributions for projected future (2040)

land-use conditions (Supplementary Information). The third set-up, global priorities (restrictive 2040), uses effective distributions that were calculated from projected future (2040) land-use conditions following stronger or more negative impacts of land use on species and ecosystems (Supplementary Information). We also analysed the potential effect that projected land-use change could have on priorities for threatened species. To this end, we defined two additional set-ups: global priorities for threatened species (present time), and global priorities for threatened species (2040) (Extended Data Table 2). In both, only threatened species (extinction risk categories vulnerable, endangered and critically endangered) are assigned standard weights as described above, whereas ecoregions and all other species are not included in the prioritization.

National analyses. To analyse the influence of national planning as opposed to globally coordinated planning^{21,72–75}, we used additional methods that produce country-specific priorities on the basis of the ranges of species and extents of ecoregions exclusively within the country boundaries^{21,53,76}. A similar approach has been used previously, at a much coarser resolution, to reveal a severe loss of performance and the emergence of edge artefacts in national conservation planning when compared to continentally coordinated planning²¹. However, the present analysis addressed a different problem: the expansion of the current global PA network, considering the effects of land-use on species distributions for present and projected future (2040) conditions. Two prioritization set-ups were defined to investigate national priorities: national priorities (present time) and national priorities (2040) (Extended Data Table 2). In both, national priorities were developed for every country considering separately the distributions of all ecoregions and species occurring in each of them, using the strong administrative priorities analysis type⁷⁶, delimited by the national boundaries derived from the GADM.

Interpreting and comparing analyses. Results were compared statistically, spatially and against well-known regional-scale global priority maps, such as the map of biodiversity hotspots revisited, 2011 revision^{13,77–79}, and the centres of plant diversity⁸⁰. The plots and statistics provided for small-range species concern those species with range size smaller than $50,000 \text{ km}^2$. In the figures and Supplementary Information, all the box plots include median, twenty-fifth and seventy-fifth percentiles (boxes), whiskers and outliers. The whiskers are extremes that are 1.5 times the height of the boxes (or interquartile range) above or below the boxes.

The maps presented here have been generated as continuous rankings of the whole land surface of the Earth. The spatial priorities resulting from our analyses are continuous estimations of the importance of the contribution of cells or sites to the global PA network. These data should not be interpreted as if they prescribed hard thresholds or decisions. Also, robust decision-making requires careful consideration of the different types of uncertainties that necessarily affect such priority maps. Two intertwined issues that are further analysed below deserve special attention: effective spatial resolution and omission and commission errors.

Maps of uncertainty corresponding to our main results are provided in Supplementary Information, showing that the spatial location of priorities is fairly consistent even when as much as 33% of additional, simulated, commission error is introduced into the species' distribution data. The degree of uncertainty in the ranking of sites or cells is considerably higher in national priorities as compared to global priorities, especially around borders of countries with edge effects.

Comparison with KBAs and other site-scale prioritizations. To test the reliability and usefulness of the priority ranking maps presented here when considering additional taxonomic groups, we compared these maps with important sites for biodiversity conservation. We compared our results with KBAs^{81,82}. These sites have been identified as the result of processes that follow an essentially different methodology and are based on partially different data, with better access to local expertise and sources of information. We analysed three national lists of KBAs: Madagascar, Myanmar and the Philippines⁸³. The list of KBAs of the Philippines^{84,85} contains 284 sites (151 terrestrial), ranging from 8 to $339,000 \text{ ha}$ of area. The KBAs of Myanmar⁸⁶ retrieved from Myanmar Biodiversity (<http://www.myanmarbiodiversity.org>) are a total of 132 sites of size ranging from 0.4 to $11,300 \text{ km}^2$. In Madagascar, a total of 1,218 sites of high or potential interest for conservation have been identified, with areas ranging from <1 to $372,000 \text{ ha}$ (ref. 87), in which sites of high potential for conservation have been identified as KBAs. In all cases, we restricted our analysis to terrestrial areas. Results (see Supplementary Information) confirm that the priority ranking maps presented here would target KBAs to a large extent, effectively inverting the trend towards less representation of important sites that has been observed in recent PA network expansion. This also provides evidence that to a notable extent, the global PA network expansion areas identified here can be efficient and representative for other biodiversity not directly considered in this study^{26,88}. This comparison with important sites is an example that the high-resolution priority ranking maps presented here can help to bridge the gap between large-scale conservation planning assessments, regional scale assessments, and site-scale assessments.

Spatial resolution. The spatial resolution⁸⁹ or grain size⁴⁷ has a notable effect on the outcomes of systematic conservation planning assessments⁸⁹, and a comparison between different results obtained for different resolutions is not strictly possible. Notwithstanding

this, we analysed how our results would vary when using range maps of species and ecoregions scaled at different coarser resolutions, taking as reference the results obtained for a resolution of 0.01667° . We compared these with results obtained for different resolutions: 0.05° , 0.1° and 0.2° . Previous related studies have used species range distributions at comparable resolutions: 0.125° (ref. 43), 0.333° (ref. 44), 10×10 km, approximately equivalent to 0.1° , or even polygons and ellipses with their full resolution⁴. We analysed the correlation between the different rankings obtained as well as the overlap between the areas identified as best candidates for expansion of the global PA network to 17% of the terrestrial world (Supplementary Information). The coarser resolution priority ranking maps were compared with upscaled versions of the reference priority ranking maps, generated by calculating median values of blocks of cells. We also compared the distribution of these expansion areas by latitude bins (Supplementary Information).

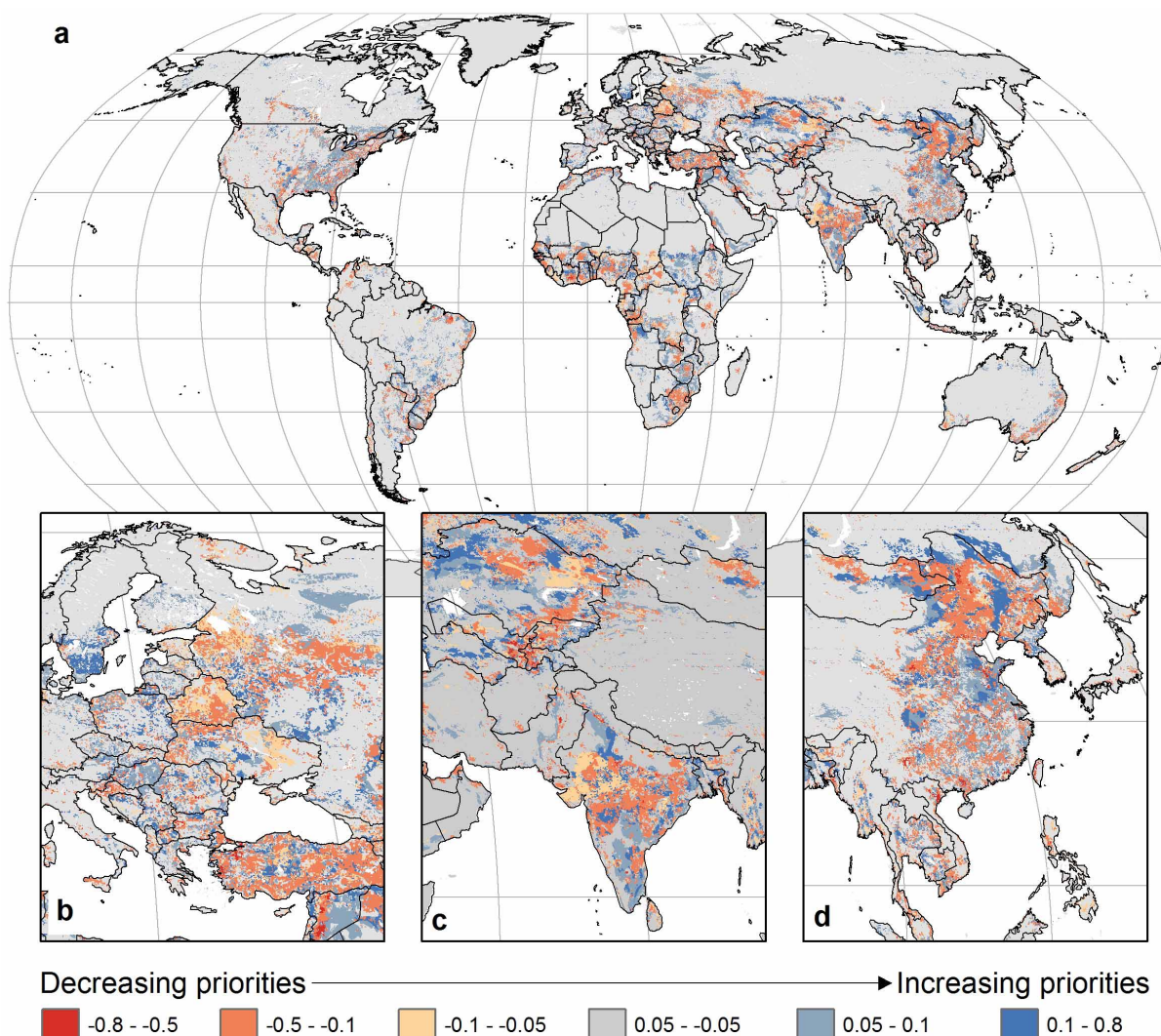
We used three measures of correlation: the Pearson correlation coefficient, the Spearman's rank correlation and the Kendall tau⁹⁰. The Pearson product-moment correlation coefficient is a measure of linear correlation between two priority rankings in this context. It takes values between -1 and $+1$, with $+1$ denoting total positive correlation. The Spearman's rank correlation coefficient is in contrast a nonparametric measure of statistical dependence that evaluates to what extent the relationship between two rankings can be described by a monotonic function. Perfect correlation of $+1$ or -1 indicates that each ranking is a perfect monotone function of the other. The Kendall tau correlation coefficient is an alternative nonparametric statistic that measures the rank correlation between two rankings, or similarity in the ordering of the rankings. We also compared aggregated results, such as the distribution of expansion areas by latitudinal bins, finding that our conclusions are robust with respect to the analysis resolution.

Omission and commission errors. There are different issues associated with different types of species occurrence data^{47,55}. In particular, different types of occurrence data, such as point localities, range maps and predicted distributions, are more or less likely to present omission and commission error. The species' range maps used here are very likely to contain important commission errors because of the nature of such maps^{37–39,47}. By contrast, omission errors can be expected to be very infrequent in these maps⁹¹. This can lead to a systematic overestimation of occurrence and representation of biodiversity in spatial prioritization. When using range maps, it is recommended to assess the sensitivity to commission errors when selecting areas for conservation in systematic conservation planning³⁷.

We performed an assessment of the sensitivity of our results to potential commission errors. We added random omissions to all the effective range maps, that is, in addition to the constraining of original range maps by land-use models, we introduced a varying percentage or rate of artificial omissions ranging from 5 to 15%, choosing coordinates and species at random. These random omissions are introduced in addition to the discounting of species' ranges by an average of 14.12% (present) and 23.99% (future, 2040) from the original range maps, reflecting human land use. We then evaluated the correlation between the different ranking maps, and the overlap between the different expansion areas obtained for different rates of artificial omission rates. These results (see Supplementary Information) give an indication of the sensitivity of our results to potential commission errors in the distribution maps. For rates of random omissions between 3.3 and 25%, the difference in average coverage of species in top 17% areas is $<2.5\%$ and the difference in expansion areas ranges between 1 and 10%. On the basis of this analysis we also generated maps of uncertainty that show, for different confidence intervals, how the ranking of top 17% areas would change owing to commission errors (Supplementary Information). This is a simple quantitative sensitivity analysis with two unrealistic assumptions that make it demanding. First, artificial omissions are generated randomly, producing a scattered cloud of omissions in addition to a discounting pattern that reflects human land use, whereas real commission errors can be expected to follow a non-random pattern. Second, we use the same rate of randomly introduced commission errors for all species (while larger range species tend to have lower rates of commission errors^{39,55}). Also, the uncertainty maps presented in the Supplementary Information were generated for the highest rate of commission error introduced into the species' distribution data (33%).

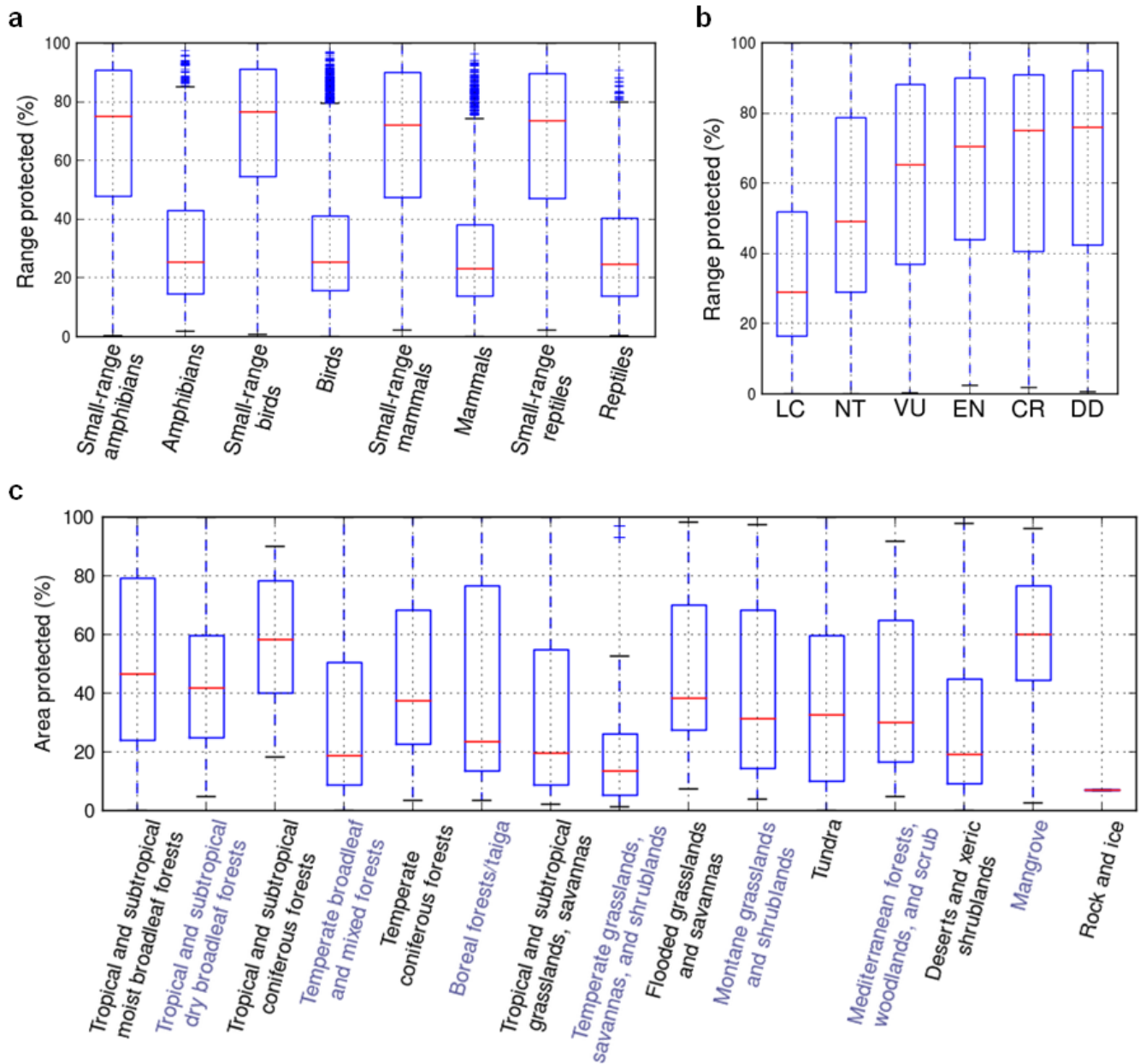
31. Wessel, P. & Smith, W. H. F. A global, self-consistent, hierarchical, high-resolution shoreline database. *J. Geophys. Res. Solid Earth* **101**, 8741–8743 (1996).
32. BirdLife International and NatureServe. *Bird Species Distribution Maps of the World*. Version 3.0 <http://www.birdlife.org/datazone/info/spcdownload> (2013).
33. BirdLife International. *BirdLife's Global Species Programme*. <http://www.birdlife.org/datazone/species> (2013).
34. Stuart, S. N. *et al.* Status and trends of amphibian declines and extinctions worldwide. *Science* **306**, 1783–1786 (2004).
35. Schipper, J. *et al.* The status of the world's land and marine mammals: diversity, threat, and knowledge. *Science* **322**, 225–230 (2008).
36. Gaston, K. J. & Fuller, R. A. The sizes of species' geographic ranges. *J. Appl. Ecol.* **46**, 1–9 (2009).
37. Rondinini, C., Wilson, K. A., Boitani, L., Grantham, H. & Possingham, H. P. Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecol. Lett.* **9**, 1136–1145 (2006).
38. Hurlbert, A. H. & Jetz, W. Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proc. Natl Acad. Sci. USA* **104**, 13384–13389 (2007).
39. Jetz, W., Sekercioglu, C. H. & Watson, J. E. M. Ecological correlates and conservation implications of overestimating species geographic ranges. *Conserv. Biol.* **22**, 110–119 (2008).
40. Cantú-Salazar, L. & Gaston, K. J. Species richness and representation in protected areas of the Western hemisphere: discrepancies between checklists and range maps. *Divers. Distrib.* **19**, 782–793 (2013).
41. Rodrigues, A. S. L. *et al.* Global gap analysis: priority regions for expanding the global protected-area network. *Bioscience* **54**, 1092–1100 (2004).
42. Jenkins, C. N., Pimm, S. L. & Joppa, L. N. Global patterns of terrestrial vertebrate diversity and conservation. *Proc. Natl Acad. Sci. USA* **110**, E2602–E2610 (2013).
43. Strassburg, B. B. N. *et al.* Impacts of incentives to reduce emissions from deforestation on global species extinctions. *Nature Clim. Change* **2**, 350–355 (2012).
44. Venter, O. *et al.* Targeting global protected area expansion for imperiled biodiversity. *PLoS Biol.* **12**, e1001891 (2014).
45. Böhm, M. *et al.* The conservation status of the world's reptiles. *Biol. Conserv.* **157**, 372–385 (2013).
46. Somveille, M., Manica, A., Butchart, S. H. M. & Rodrigues, A. S. L. Mapping global diversity patterns for migratory birds. *PLoS ONE* **8**, e70907 (2013).
47. Rahbek, C. The role of spatial scale and the perception of large-scale species-richness patterns: scale and species-richness patterns. *Ecol. Lett.* **8**, 224–239 (2005).
48. Schulman, L., Toivonen, T. & Ruokolainen, K. Analysing botanical collecting effort in Amazonia and correcting for it in species range estimation: Amazonian collecting and range estimation. *J. Biogeogr.* **34**, 1388–1399 (2007).
49. OECD. *OECD Environmental Outlook to 2050: The Consequences of Inaction*. <http://www.oecd.org/environment/oecdenvironmentaloutlookto2050theconsequencesofinaction.htm> (2012).
50. Balmford, A., Green, R. & Phalan, B. What conservationists need to know about farming. *Proc. R. Soc. B* **279**, 2714–2724 (2012).
51. Laurance, W. F., Sayer, J. & Cassman, K. G. Agricultural expansion and its impacts on tropical nature. *Trends Ecol. Evol.* **29**, 107–116 (2014).
52. Leathwick, J. R., Moilanen, A., Ferrier, S. & Julian, K. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biol. Conserv.* **143**, 984–991 (2010).
53. Moilanen, A. *et al.* *Zonation — Spatial Conservation Planning Methods and Software Version 4*, User Manual (Univ. Helsinki, 2014).
54. Mendenhall, C. D., Karp, D. S., Meyer, C. F. J., Hadly, E. A. & Daily, G. C. Predicting biodiversity change and averting collapse in agricultural landscapes. *Nature* **509**, 213–217 (2014).
55. Rondinini, C. *et al.* Global habitat suitability models of terrestrial mammals. *Philos. Trans. R. Soc. B* **366**, 2633–2641 (2011).
56. Salafsky, N. *et al.* A standard lexicon for biodiversity conservation: unified classifications of threats and actions. *Conserv. Biol.* **22**, 897–911 (2008).
57. Rodrigues, A. S. L. Improving coarse species distribution data for conservation planning in biodiversity-rich, data-poor, regions: no easy shortcuts. *Anim. Conserv.* **14**, 108–110 (2011).
58. Arponen, A., Heikkinen, R. K., Thomas, C. D. & Moilanen, A. The value of biodiversity in reserve selection: representation, species weighting, and benefit functions. *Conserv. Biol.* **19**, 2009–2014 (2005).
59. Lehtomäki, J. & Moilanen, A. Methods and workflow for spatial conservation prioritization using Zonation. *Environ. Model. Softw.* **47**, 128–137 (2013).
60. Lehtomäki, J., Tomppo, E., Kuokkanen, P., Hanski, I. & Moilanen, A. Applying spatial conservation prioritization software and high-resolution GIS data to a national-scale study in forest conservation. *For. Ecol. Manage.* **258**, 2439–2449 (2009).
61. Margules, C. & Sarkar, S. *Systematic Conservation Planning* (Cambridge Univ. Press, 2007).
62. Kukkala, A. S. & Moilanen, A. Core concepts of spatial prioritisation in systematic conservation planning. *Biol. Rev. Camb. Philos. Soc.* **88**, 443–464 (2013).
63. Moilanen, A. Landscape zonation, benefit functions and target-based planning: unifying reserve selection strategies. *Biol. Conserv.* **134**, 571–579 (2007).
64. Laitila, J. & Moilanen, A. Use of many low-level conservation targets reduces high-level conservation performance. *Ecol. Modell.* **247**, 40–47 (2012).
65. Butchart, S. H. M. *et al.* Measuring global trends in the status of biodiversity: red list indices for birds. *PLoS Biol.* **2**, e383 (2004).
66. Balmford, A., Gaston, K. J., Blyth, S., James, A. & Kapos, V. Global variation in terrestrial conservation costs, conservation benefits, and unmet conservation needs. *Proc. Natl Acad. Sci. USA* **100**, 1046–1050 (2003).
67. Moore, J., Balmford, A., Allnutt, T. & Burgess, N. Integrating costs into conservation planning across Africa. *Biol. Conserv.* **117**, 343–350 (2004).
68. Naidoo, R. & Iwamura, T. Global-scale mapping of economic benefits from agricultural lands: Implications for conservation priorities. *Biol. Conserv.* **140**, 40–49 (2007).
69. Eklund, J., Arponen, A., Visconti, P. & Cabeza, M. Governance factors in the identification of global conservation priorities for mammals. *Philos. Trans. R. Soc. B* **366**, 2661–2669 (2011).
70. Waldron, A. *et al.* Targeting global conservation funding to limit immediate biodiversity declines. *Proc. Natl Acad. Sci. USA* **110**, 12144–12148 (2013).
71. Vincent, J. R. *et al.* Tropical countries may be willing to pay more to protect their forests. *Proc. Natl Acad. Sci. USA* **111**, 10113–10118 (2014).

72. Rodrigues, A. S. L. & Gaston, K. J. Rarity and conservation planning across geopolitical units. *Conserv. Biol.* **16**, 674–682 (2002).
73. Kark, S., Levin, N., Grantham, H. S. & Possingham, H. P. Between-country collaboration and consideration of costs increase conservation planning efficiency in the Mediterranean Basin. *Proc. Natl Acad. Sci. USA* **106**, 15368–15373 (2009).
74. Gustafsson, L. *et al.* Natural versus national boundaries: the importance of considering biogeographical patterns in forest conservation policy. *Conserv. Lett.* <http://dx.doi.org/10.1111/conl.12087> (13 February 2014).
75. Mazar, T., Possingham, H. P. & Kark, S. Collaboration among countries in marine conservation can achieve substantial efficiencies. *Divers. Distrib.* **19**, 1380–1393 (2013).
76. Moilanen, A. & Arponen, A. Administrative regions in conservation: balancing local priorities with regional to global preferences in spatial planning. *Biol. Conserv.* **144**, 1719–1725 (2011).
77. Mittermeier, R. A. *et al.* *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions* (Conservation International, 2005).
78. Potapov, P. *et al.* Mapping the world's intact forest landscapes by remote sensing. *Ecol. Soc.* **13**, 51 (2008).
79. Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858 (2000).
80. Davis, S. D., Heywood, V. H. & Hamilton, A. C. *Centres of Plant Diversity*. <http://www.unep-wcmc.org/resources-and-data> (2013).
81. Eken, G. *et al.* Key biodiversity areas as site conservation targets. *Bioscience* **54**, 1110–1118 (2004).
82. Knight, A. T. *et al.* Improving the key biodiversity areas approach for effective conservation planning. *Bioscience* **57**, 256–261 (2007).
83. BirdLife International, Conservation International, IUCN & UNEP-WCMC. *Protected Area and Key Biodiversity Area*. Data downloaded from the Integrated Biodiversity Assessment Tool (IBAT) <https://www.ibat-alliance.org/ibat-conservation> (2014).
84. Ambal, R. G. R. *et al.* Key biodiversity areas in the Philippines: priorities for conservation. *J. Threat. Taxa* **4**, 2788–2796 (2012).
85. Conservation International. *Priority Sites for Conservation in the Philippines: Key Biodiversity Areas*. http://www.conservation.org/global/philippines/publications/Documents/KBA_Booklet.pdf (2006).
86. Tordoff, A. W., Baltzer, M. C., Fellowes, J. R., Pilgrim, J. D. & Langhammer, P. F. Key biodiversity areas in the Indo-Burma Hotspot: process, progress and future directions. *J. Threat. Taxa* **4**, 2779–2787 (2012).
87. SAPM. Les sites du Système des Aires Protégées de Madagascar. Shapefile des sites du SAPM. Arrêté Interministériel n°9874/2013 Modifiant Certaines Dispositions de l'Arrêté n°52005/2010. <http://atlas.rebioma.net/> (2014).
88. Jetz, W., McPherson, J. M. & Guralnick, R. P. Integrating biodiversity distribution knowledge: toward a global map of life. *Trends Ecol. Evol.* **27**, 151–159 (2012).
89. Arponen, A., Lehtomäki, J., Leppänen, J., Tomppo, E. & Moilanen, A. Effects of connectivity and spatial resolution of analyses on conservation prioritization across large extents. *Conserv. Biol.* **26**, 294–304 (2012).
90. Myers, J. L., Well, A. & Lorch, R. F. *Research Design and Statistical Analysis*. (Routledge, 2010).
91. Ficetola, G. F. *et al.* An evaluation of the robustness of global amphibian range maps. *J. Biogeogr.* **41**, 211–221 (2014).



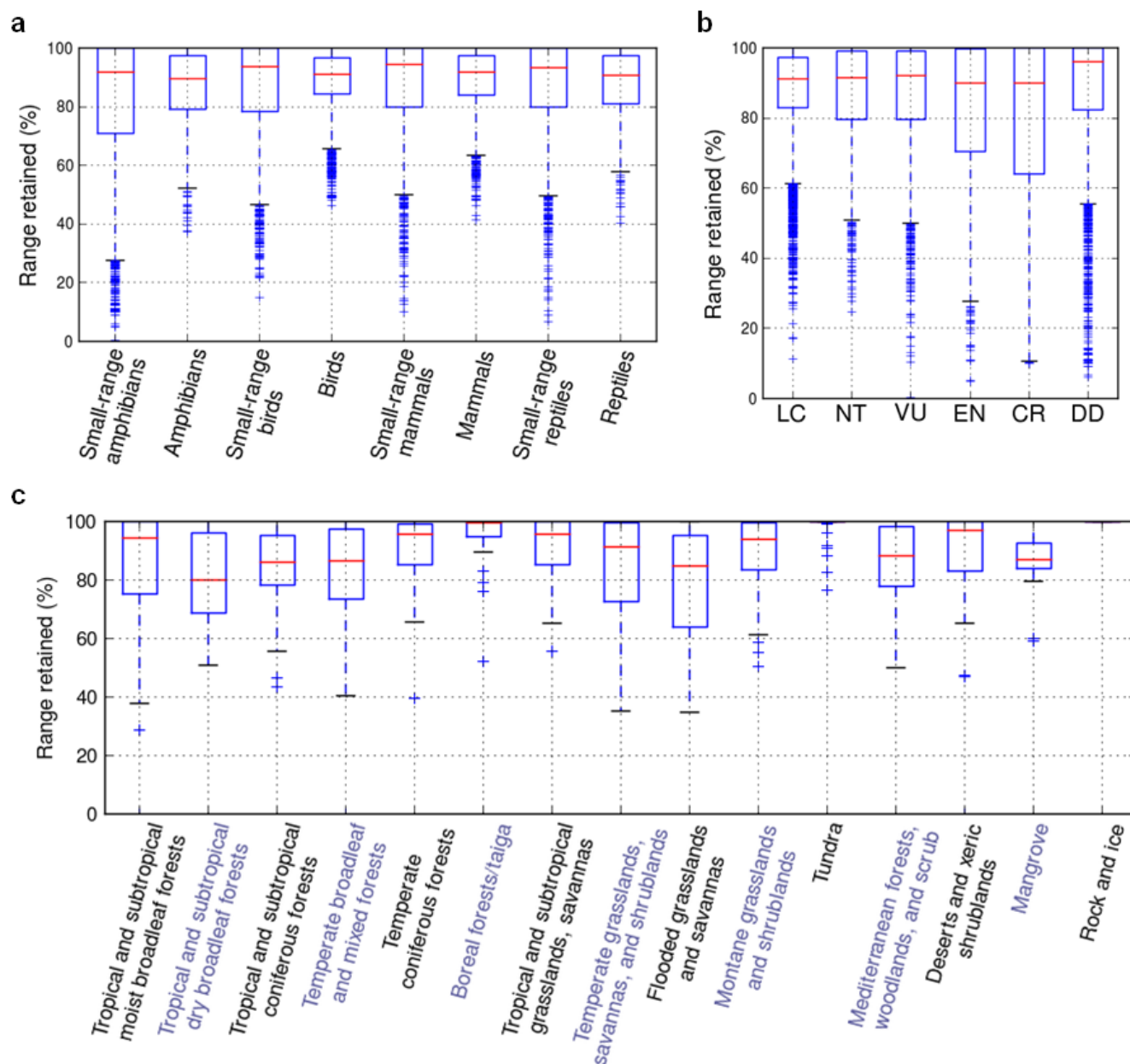
Extended Data Figure 1 | Changes in spatial conservation priority between present and future (2040). a–d, The top areas for PA expansion remain relatively stable: the congruence between priority expansion areas for present and projected future land use is 77.9%. Despite relatively high congruence

(Supplementary Information), there are important localized differences. The biggest declines in priority would happen in China (d), India (c), eastern Europe and Turkey (b), whereas the changes are more subtle in sub-Saharan Africa and the Americas.



Extended Data Figure 2 | Box plots of protection of effective range (species) and effective extent (ecoregions) in the expanded global PA system, under projected future (2040) land-use conditions. a, b, Summaries of coverage for species grouped by taxonomic groups (classes) (a) and IUCN status (b). c, Ecoregions grouped by biome. These box plots show median values,

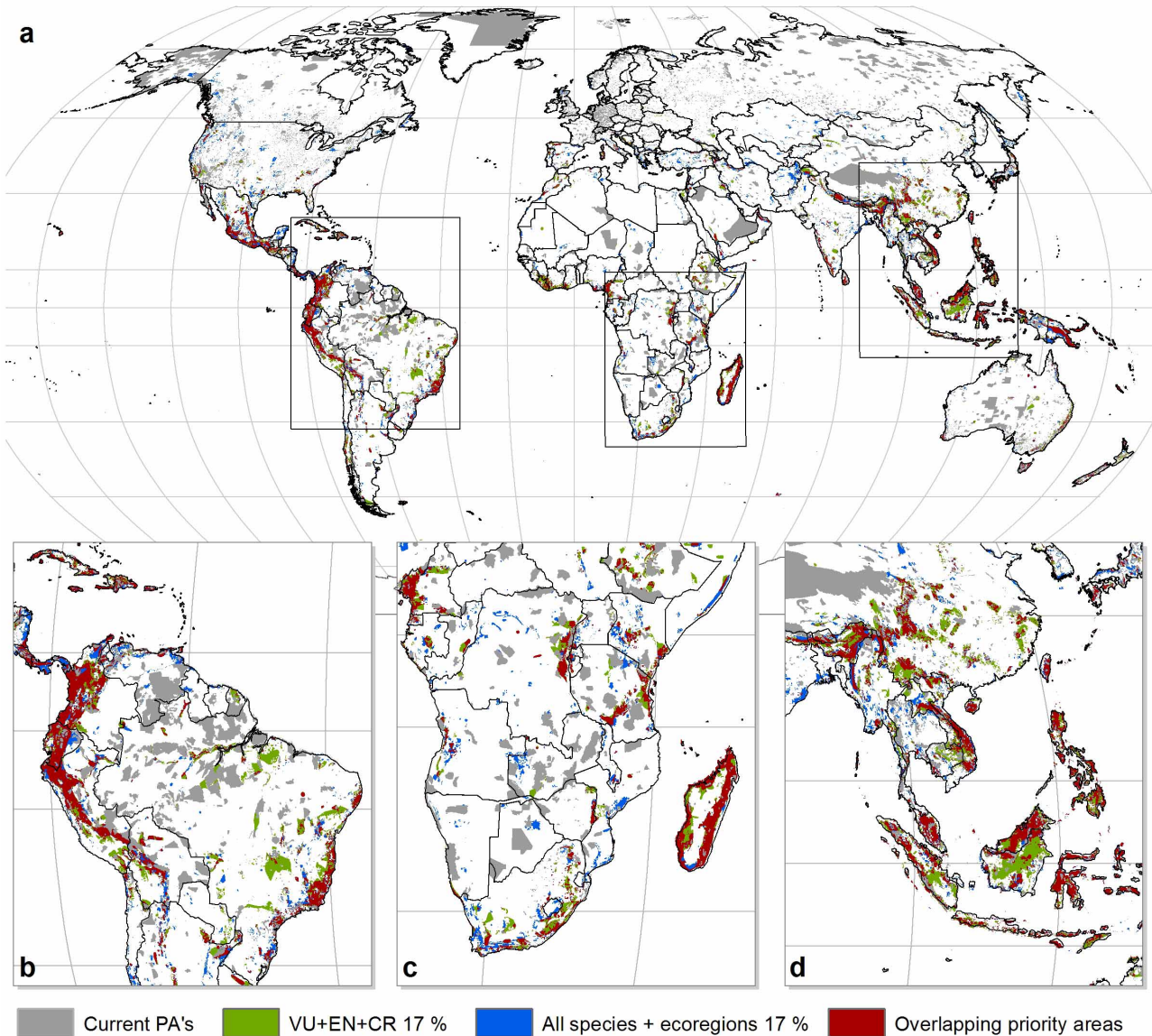
twenty-fifth and seventy-fifth percentiles (boxes), whiskers (1.5 times the interquartile range) and outliers. Protection levels are well balanced for different species groups, and between species and ecoregions. Protection levels tend to be lower for less threatened species, as these tend to have wider ranges.



Extended Data Figure 3 | Box plots of loss of effective range (species) and effective extent (ecoregions) from projected land-use changes by 2040.

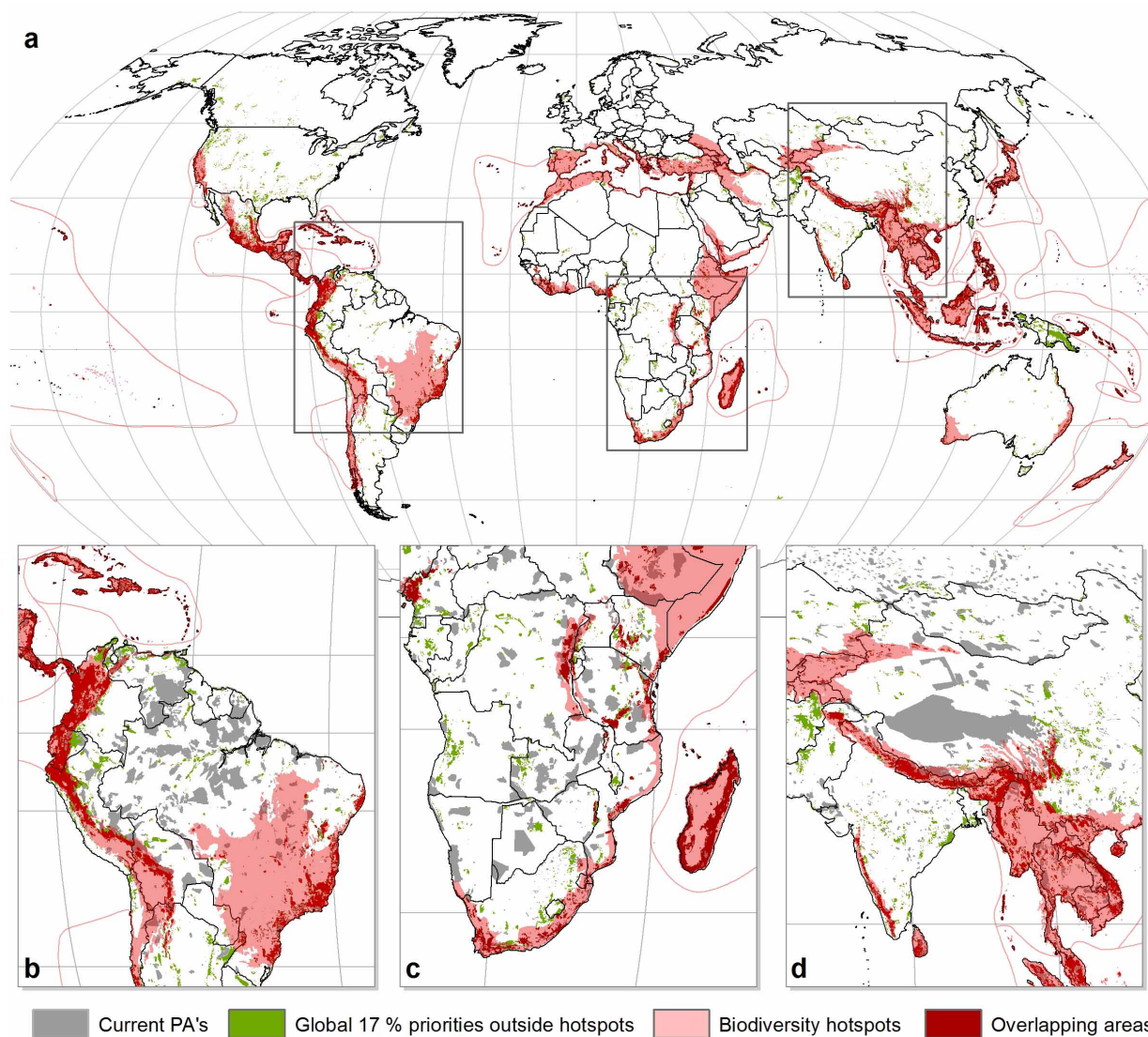
a, Species grouped by taxonomic groups (classes), distinguishing small-range species (range size $<50,000 \text{ km}^2$). **b**, Species grouped by IUCN threat status.

c, Ecoregions grouped by biome. The proportion of species that are expected to lose a significant fraction of their habitat is higher for species with a higher threat status.

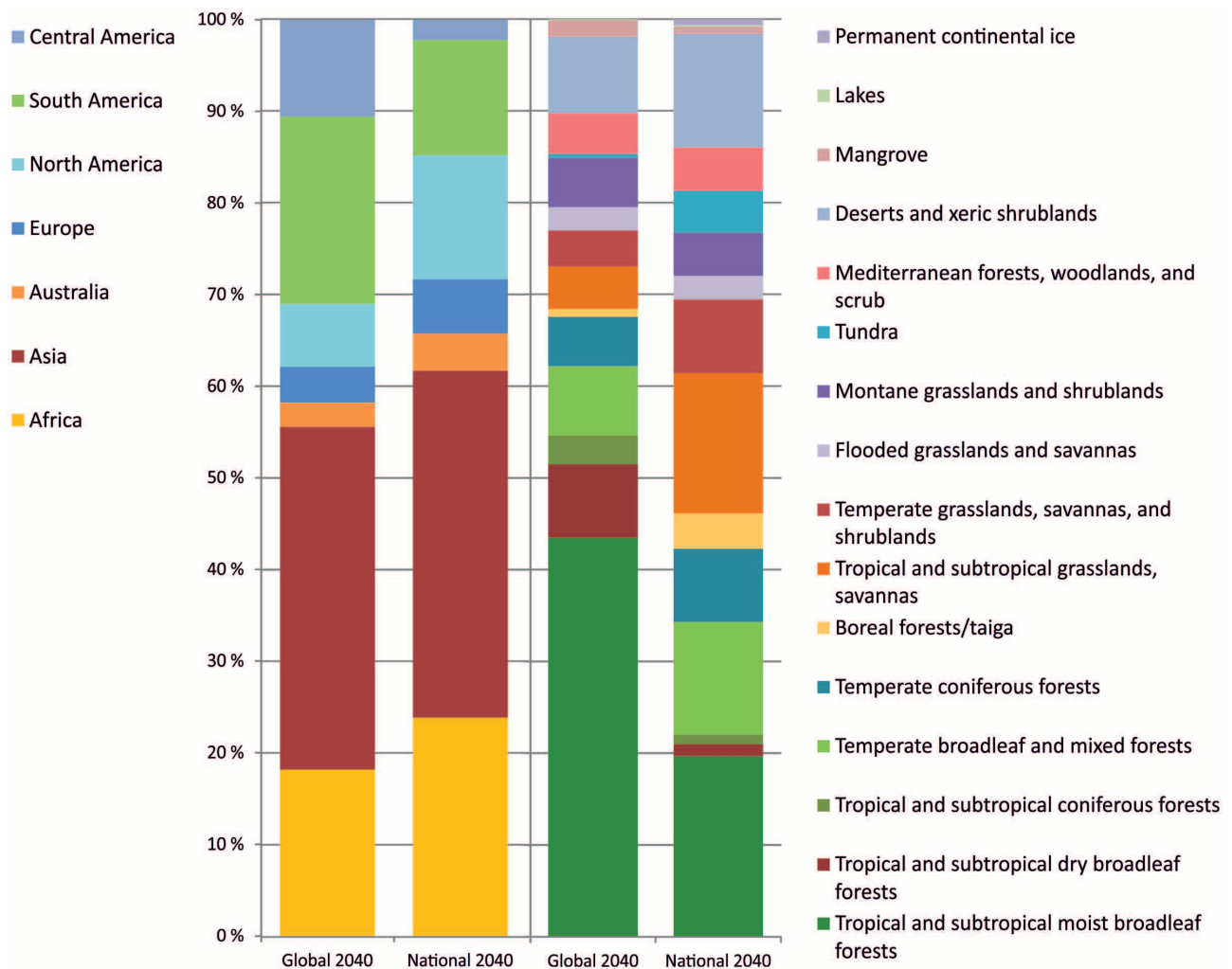


Extended Data Figure 4 | Comparison of priority areas for threatened species, and all species and ecoregions, both considering projected future land-use (2040). a–d, The overall overlap of the respective top 17% priority areas is 62%. Priorities are highly congruent in most biodiversity hotspots of the

world. More top priority areas are identified for threatened species in the tropics, whereas there are more top priority areas in higher latitudes for ecoregions and all vertebrate species. IUCN threat categories: critically endangered (CR), endangered (EN) and vulnerable (VU).



Extended Data Figure 5 | Global expansion priority areas for projected future (2040) land-use. a–d, Some of the areas in which the largest spatially contiguous overlaps occur are highlighted. Areas that overlap with biodiversity hotspots (full red) and those outside hotspots (green) are shown.



Extended Data Figure 6 | Stacked bar plot showing the distributions of 17% expansion areas across different continents (left) and biomes (right), for future (2040) land-use. When following national priorities, the distribution of expansion areas tends to be more balanced between biomes, at the expense of lower average protection of species and ecoregions, particularly favouring

grasslands over tropical forests. The continental responsibility for Asia is virtually independent on whether national or global priorities are followed, whereas if planning is made nationally, responsibility clearly increases in Africa and North America and decrease in Central and South America. These patterns are stable across time (Supplementary Information).

Extended Data Table 1 | Species with projected effective range loss above 30, 50 and 70% for land-use change projected for 2040

Threat status	Number of species	Species with >30% effective range loss	Species with >50% effective range loss	Species with >70% effective range loss
LC	14978	953	192	11
NT	1783	234	51	4
VU	2092	320	103	14
EN	1804	443	180	33
CR	986	277	161	66
DD	2954	439	207	93

Species are grouped by their category of extinction risk on the IUCN red list of threatened species. The values shown reflect changes in the effective range of species as a consequence of projected future (2040) land-use intensification.

Extended Data Table 2 | Summary of protection levels of species ranges and ecoregions area for the expanded (17%) PA system

	global priorities (present time)	global priorities (2040)	global priorities (pessimistic 2040)	global priorities for threatened species (present time)	global priorities for threatened species (2040)	national priorities (present time)	national priorities (2040)
LC	46.8	42.5	42.0	45.7	41.2	30.0	27.7
NT	68.7	61.3	59.6	67.3	60.2	42.5	38.6
VU	80.7	70.8	68.4	86.4	75.4	54.9	48.9
EN	92.1	77.3	72.6	94.7	79.7	73.6	62.9
CR	95.8	77.4	78.3	97.2	78.8	85.4	69.5
DD	86.2	77.3	76.3	76.0	69.3	71.3	64.7
All non-threatened vertebrates	54.7	49.4	48.6	51.8	46.7	37.3	34.2
All threatened vertebrates	87.9	74.5	71.7	91.6	77.6	67.8	58.1
All vertebrates	61.2	54.3	53.0	59.7	52.9	43.3	38.9
All ecoregions	55.4	48.9	47.4	42.8	38.2	36.6	33.8

Protection levels are reported as average percentages of the (effective) global range size (species) or area (ecoregions), covered by 17% top priority areas for present and projected future (2040) land-use conditions.

Signatures of aestivation and migration in Sahelian malaria mosquito populations

A. Dao¹, A. S. Yaro¹, M. Diallo¹, S. Timbiné¹, D. L. Huestis², Y. Kassogue¹, A. I. Traoré¹, Z. L. Sanogo¹, D. Samaké¹ & T. Lehmann²

During the long Sahelian dry season, mosquito vectors of malaria are expected to perish when no larval sites are available; yet, days after the first rains, mosquitoes reappear in large numbers. How these vectors persist over the 3–6-month long dry season has not been resolved, despite extensive research for over a century^{1–3}. Hypotheses for vector persistence include dry-season diapause (aestivation) and long-distance migration (LDM); both are facets of vector biology that have been highly controversial owing to lack of concrete evidence. Here we show that certain species persist by a form of aestivation, while others engage in LDM. Using time-series analyses, the seasonal cycles of *Anopheles coluzzii*, *Anopheles gambiae sensu stricto* (s.s.), and *Anopheles arabiensis* were estimated, and their effects were found to be significant, stable and highly species-specific. Contrary to all expectations, the most complex dynamics occurred during the dry season, when the density of *A. coluzzii* fluctuated markedly, peaking when migration would seem highly unlikely, whereas *A. gambiae* s.s. was undetected. The population growth of *A. coluzzii* followed the first rains closely, consistent with aestivation, whereas the growth phase of both *A. gambiae* s.s. and *A. arabiensis* lagged by two months. Such a delay is incompatible with local persistence, but fits LDM. Surviving the long dry season *in situ* allows *A. coluzzii* to predominate and form the primary force of malaria transmission. Our results reveal profound ecological divergence between *A. coluzzii* and *A. gambiae* s.s., whose standing as distinct species has been challenged, and suggest that climate is one of the selective pressures that led to their speciation. Incorporating vector dormancy and LDM is key to predicting shifts in the range of malaria due to global climate change⁴, and to the elimination of malaria from Africa.

Over half a million malarial deaths still occur annually, mostly in sub-Saharan Africa⁵. Transmitted by *Anopheles gambiae* s.s., *A. coluzzii* (previously known as the *A. gambiae* S and M molecular forms⁶), *A. arabiensis* and *A. funestus*, malaria is widespread, including in dry savannahs and semi-arid areas. Persistence of malaria in areas where the surface waters required for larval development are absent for several months a year^{2,7–12} has been the subject of much interest, as it has long been recognized that, during the dry season, reproductively quiescent adult mosquitoes are ideally suited for vector control^{13,14}. Recent findings suggested that aestivation is used by *A. coluzzii* to persist throughout the dry season^{14–18}; yet, more definitive evidence is required to fully resolve this question.

Data from a five-year study of Sahelian *A. coluzzii*, *A. gambiae* s.s. and *A. arabiensis* population densities at an unparalleled resolution were subjected to time-series analyses to isolate the seasonal components, assess their magnitude, and determine if they were stable or time-varying (Methods). This statistical framework allowed identification of salient elements of the seasonal cycle of each species, providing unique ecological signatures, which were then deciphered to determine if populations endured the dry season locally or if populations recolonized the area by migration.

From September 2008 to August 2013, a total of 40,195 *A. gambiae sensu lato* (s.l.) (28,547 females and 11,648 males) were collected in the

Sahelian village of Thierola, Mali, during 511 collection days (Fig. 1 and Extended Data Fig. 1; Extended Data Table 1, Supplementary Information). The complexity of the population dynamics of *A. gambiae* s.l. was epitomized by dramatic fluctuations during the dry season (Extended Data Figs 2 and 3). Putative seasonal elements were visually identified (Methods and Extended Data Table 2), providing a descriptive framework and expectations, to aid the interpretation of the statistical results. Briefly, the population growth phase (June–August) started ~3 weeks after the first rain, resulting in the wet-season peak (September–October). Density declined as larval sites dried (November), reaching its dry-season minima in February–March. Surprisingly, density started rising halfway into the dry season (March) and culminated in a dramatic dry-season peak lasting <7 days, returning to the typical low density weeks later (April–May), and ending with the first rain surge, 3–7 days after the first rains (Extended Data Fig. 3 and Extended Data Table 2).

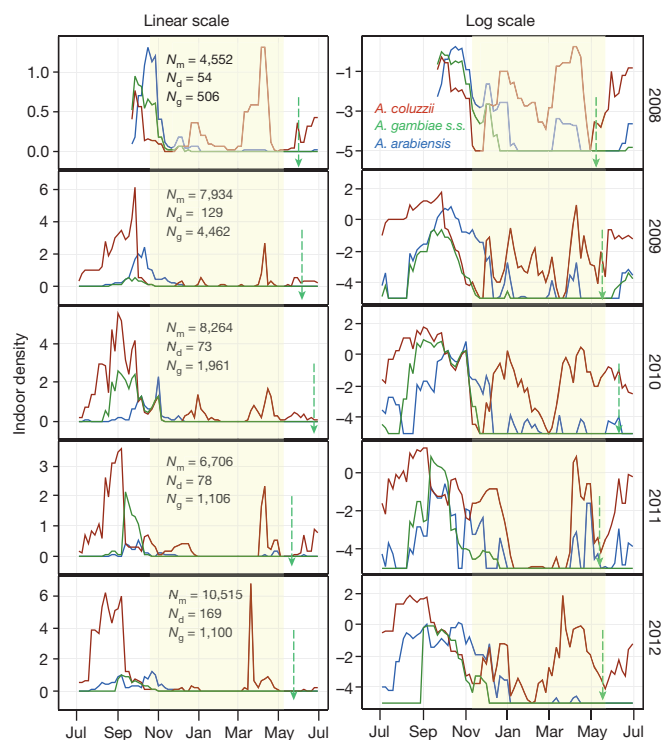


Figure 1 | Species-specific population dynamics of the members of *Anopheles gambiae* s.l. Average densities of *Anopheles coluzzii* (red), *A. gambiae* s.s. (green), and *A. arabiensis* (blue) are shown on linear and natural logarithm scales from July to June of every year, portraying changes both at low and high density ranges. Green arrows mark the first rain and tan background denotes the dry season. N_m , N_d , and N_g denote sample size of *A. gambiae* s.l., the number of collection days, and the number genotyped to species, respectively (Methods). Foreground shading indicates a gap in sampling (December–March 2008) when imputed values were used (Methods).

¹International Center for Excellence in Research (ICER), University of Sciences, Techniques and Technologies, BP 1805, Bamako, Mali. ²Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, Maryland 20852, USA.

Table 1 | Unobserved component time-series (final) models of the population dynamics for each taxon (Methods)

Taxon*	Parameter	Variance (stochastic)†	P (variance)†	Deterministic effects estimated‡	P (effect)‡
<i>A. gambiae s.l.</i> $R^2 = 0.76$ AIC = 793.5 BIC = 808.1	Seasonals§	0 (fixed)	NA	See Fig. 2	0.0001
	Level E var	0 (fixed)	NA	–1.23	0.0001
	Cycle¶ DampF	0.7	0.0001		0.99
	Cycle¶ period	5,758.7	0.99		
	Cycle¶ E var	0.56	0.0001		
	Irreg. E var	0.0112	0.85	NA	0.95
<i>A. coluzzii</i> $R^2 = 0.72$ AIC = 844 BIC = 858	Seasonals	0 (fixed)	NA	See Fig. 2	0.0001
	Level E var	0 (fixed)	NA	–1.68	0.0001
	Cycle DampF	0.7	0.0001		0.99
	Cycle period	47.1	0.5		
	Cycle E var	0.65	0.0001		
	Irreg. E var	0.011	0.92	NA	0.95
<i>A. gambiae s.s.</i> $R^2 = 0.89$ AIC = 515 BIC = 545	Seasonals	0 (fixed)	Na	See Fig. 2	0.0001
	Level E var	0 (fixed)	Na	–4.1	0.0001
	Cycle DampF	0.76	0.0001		0.0005
	Cycle period	16.1	0.0001		
	Cycle E var	0.14	0.0017		
	Cycle2 DampF	1	0.0001		0.21
	Cycle2 period	41.1	0.0001		
	Cycle2 E var	0.0001	0.54		
	Irreg. E var	0.042	0.28		0.0003
	Irreg. AR(1)#	0.94	0.0001		
<i>A. arabiensis</i> $R^2 = 0.77$ AIC = 742 BIC = 757	Seasonals	0 (fixed)	NA	See Fig. 2	0.0001
	Level E var	0 (fixed)	NA	–3.52	0.0001
	Cycle DampF	0.69	0.0001		0.08
	Cycle period	11,966	0.99		
	Cycle E var	0.48	0.0001		
	Irreg. E var	0.00001	0.99	NA	0.99

* All models (species) include 362 observations (5-day means from 22 September 2008 and 1 September 2013, based on all *A. gambiae s.l.* and those genotyped, see Methods and Extended Data Fig. 1).

† Stochastic variance and test of significance (P (variance)) indicate whether the parameter is time varying.

‡ Effect size and test of significance (P (effect)) measure the overall deterministic effects.

§ Seasonal component was modelled by 73 dummy variables. Individual effect of each of these parameters and 95% CI are shown in Fig. 2 (see text and Methods).

|| Level is equivalent to intercept (in unobserved component model (UCM) framework, if time-varying, it results in a 'random walk' between successive time points), and was found to be fixed in all analyses.

¶ Non-seasonal stochastic (trigonometric) cycles, each defined by three parameters: a period (time difference between two successive peaks; here in units of 5-day intervals), cycle damping factor (decay in amplitude between cycles over time), and the variance of the period (Methods and Supplementary Text).

One-lag autoregressive (AR1) parameter was modelled as part of the irregular component of *A. gambiae s.s.*

E, error; Irreg., irregular; var, variance.

Time-series analysis of the log-transformed density (Extended Data Fig. 2), using an unobserved components model (Methods), was fitted for *A. gambiae s.l.* (Table 1). The model selected had a fixed level (equivalent to intercept) and no slope (trend), reflecting a stable mosquito density over the study. An additional non-seasonal cycle with a long period was also included (Methods and Supplementary Information). The variance of the seasonal component was insignificant, indicating it was not time-varying; thus, it was modelled as a fixed component, simplifying its interpretation. The seasonal component of *A. gambiae s.l.* population dynamics was highly significant ($P < 0.0001$, Table 1). The estimated seasonal variation (Fig. 2a) revealed a large gap between the 95% confidence intervals (CIs) of the wet-season peak and that of the mid-dry-season low; thus, these elements and the decline between them are statistically well-supported. Likewise, large gaps were found between the 95% CIs of the mid-dry-season low and the late-dry-season peak, between this peak, the end-dry-season low, and the following wet-season peak, indicating that these elements (and the transitional phases connecting them) were statistically supported. Other putative elements (Extended Data Table 2) had insufficient statistical support.

The putative elements of each species' seasonal cycle were identified (Extended Data Table 2). The seasonal component of all species was fixed (its variance was insignificant) and was highly significant ($P < 0.0001$, Table 1). The time-series model selected for *A. coluzzii* was structurally similar to that of *A. gambiae s.l.* (Table 1). On the basis of their 95% CIs, one wet-season peak and two dry-season peaks, which were observed in all years (Fig. 1), were statistically supported (Fig. 2b). The early wet-season decline of *A. coluzzii* produced the pre-dry-season trough in mid-November, before the last larval site dried, which was followed by an early dry-season peak in late December (Fig. 2b). Subsequently, its seasonal component was virtually identical to that of *A. gambiae s.l.*

(Fig. 2), consistent with its predominance in species composition (Extended Data Fig. 1c). The model for *A. gambiae s.s.* included two non-seasonal cycles as well as an autoregressive (lag 1) error (Table 1). Only a single peak (wet season) and a long dry-season trough were statistically discerned in *A. gambiae s.s.* (Fig. 2c). The model for *A. arabiensis* was structurally similar to that of *A. coluzzii* (Table 1). A single wet-season peak and the long dry-season trough were supported (Fig. 2d), whereas changes during the dry season were not distinguished from noise.

The species-specific signatures manifested by their population dynamics provide compelling evidence that *A. coluzzii* persists locally in the Sahel during the dry season, whereas *A. gambiae s.s.* recolonizes via LDM after the first rains; the evidence is less clear for *A. arabiensis*. First, *A. coluzzii* was present throughout the dry season (albeit in small numbers), whereas *A. gambiae s.s.* was undetected from January to May (Fig. 2 and Extended Data Fig. 4), consistent with previous studies^{9,14,19–22}. Second, the density of *A. coluzzii* rose dramatically (ten- to ninety-fold from their preceding phase) twice during the dry season (Fig. 2). Since these peaks preceded the first rain by at least six weeks, any potential migrant mosquitoes would likely perish before reproductive opportunities were available, given the absence of surface waters in the area (ruling out dry-season reproduction). Third, the most crucial evidence relates to the period when population growth starts with respect to the first rain. The onset of population growth can be defined as the first time when the lower 95% CI of the seasonal component is greater than the upper 95% CI during the preceding dry season's low phase (red arrows, Fig. 2). This phase started in June for *A. coluzzii* but in August for *A. gambiae s.s.* and *A. arabiensis* (Fig. 2). A delay of six to eight weeks in the onset of population growth for the latter two species corroborated our previous results in two other Sahelian villages, 10–25 km away from Thierola¹⁴. Commencing population growth shortly after the first rain fits well with

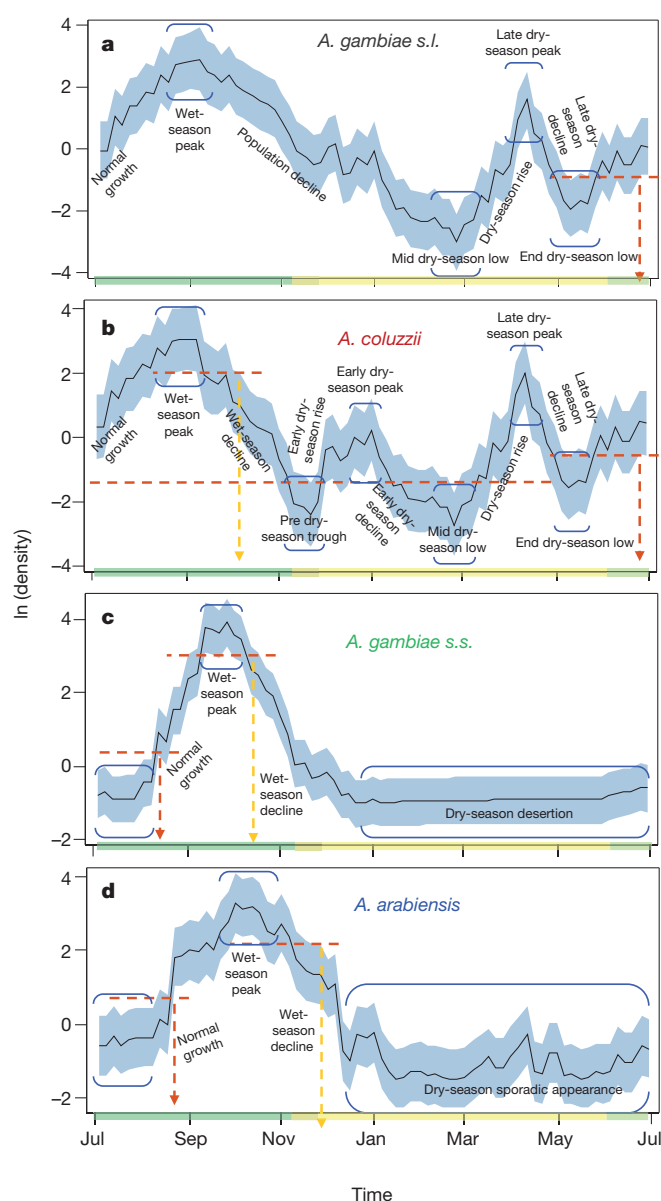


Figure 2 | Seasonal population dynamics of the members of *Anopheles gambiae* s.l. The seasonals were estimated using unobserved component time-series models (Table 1, and Methods). Bands denote 95% CI, while blue brackets surround peaks and troughs whose 95% CIs do not overlap. Red and orange arrows denote the onset and decline of population growth, respectively; defined as the earliest time when the 95% CI of the population growth (or decline) phase does not overlap with that of the preceding phase (horizontal red line). Population phase names correspond with putative elements (Extended Data Table 2). Sample sizes are based on Fig. 1.

local persistence (for example, aestivation), but a two-month ‘delay’ in that phase cannot be reconciled with it, especially contrasted with its rapid onset in *A. coluzzii*. Arrival of migrants from distant locations (and reproduction), on the other hand, may take several weeks, consistent with this delay. The earlier (August versus October) and higher wet-season peak of *A. coluzzii* is explained by the two-month ‘advantage’ it had in building its density (Fig. 2 and Extended Data Fig. 4). The prompt population growth of *A. coluzzii* is consistent with previous studies showing that its density surged over tenfold, five days after the first rain^{14,16} (egg-to-adult developmental time is ≥ 8 days) and with the recapture of one female that was captured, marked and released seven months earlier in the same village¹⁶. Fourth, density of *A. coluzzii* was declining by October, at least 4 weeks before the last larval sites dried

up, reaching its pre-dry-season trough in November, whereas *A. arabiensis* and *A. gambiae* s.s. continued to reproduce (Figs 1, 2 and Extended Data Fig. 4). This early decline in *A. coluzzii* is consistent with another hallmark of diapause, the initiation phase^{23–25}, in which insects change their behaviour and physiology and move into shelters before unfavourable conditions unfold.

During the dry season, no surface waters were available near Thierola for at least a 30 km radius, and in the distant localities where surface waters did exist, overall density was very low and *A. gambiae* s.s. was not detected until the wet season^{14,16}. The nearest high-density source is in the Niono rice cultivation area (~150 km ENE of Thierola), but it consists exclusively of *A. coluzzii*²⁶. Therefore, LDM spanning hundreds of kilometres is necessary to explain the re-colonization of *A. gambiae* s.s. Alternative explanations, including desiccation-tolerant dormant eggs, larvae or pupae, as well as larval growth in deep, underground water sources, should not be altogether dismissed, despite being contradictory to available knowledge.

Population dynamics of *A. arabiensis* exhibited mixed signatures. Statistically, it is similar to *A. gambiae* s.s., and the long delay in population growth after the rains (Figs 1, 2 and Extended Data Fig. 4) indicates that it too persists by LDM. Yet, throughout the dry season, sporadic individuals were found every year, as opposed to zero *A. gambiae* s.s. Possibly, the dominant strategy of *A. coluzzii* is expressed in a small fraction of *A. arabiensis*, consistent with previous reports of local persistence of *A. arabiensis* by aestivation in the East African Sahel²⁷ and perhaps in other parts of the West African Sahel^{10,27}. Alternatively, the occasional *A. arabiensis* recovered in the dry season could represent backcrossed hybrids between *A. coluzzii* and *A. arabiensis*.

These results provide fresh insights that dramatically change our understanding of the ecology of African malaria vectors and resolve the ‘dry-season paradox’. Ignoring the pervasive effects of dormancy and LDM limits our understanding of malaria transmission and its response to control and elimination strategies. Dormancy shapes vector composition in the Sahel, where *A. coluzzii* comprised 75% of the overall indoor vector density with its wet-season peak being at least twice as high and broad than that of either *A. gambiae* s.s. or *A. arabiensis* (Fig. 1, Supplementary Information and refs 14, 16). Arguably, dormancy underlies the heavy burden of malaria transmission in such areas by exponential amplification of human-vector cycles that culminate in intense late-wet-season transmission. Although *A. gambiae* s.s. and *A. arabiensis* predominate during the end of the wet season (October–November), we doubt that they alone can sustain the high rate of malaria transmission had *A. coluzzii* not amplified infections from June to September. Therefore, vector-control strategies that eliminate *A. coluzzii* alone may cut peak malaria transmission to very low levels. Targeting *A. coluzzii* while in its hidden shelters during the dry season is probably the most efficient control strategy, if these sites are found, but the indoor population during the early wet season and the late-dry-season peak also presents promising targets. Thus, a single residual spraying indoors in the late dry-season (for example, March) that is effective for 4 months may achieve dramatic reduction in malaria transmission in the following wet season¹⁴. Moreover, the spread of introduced genes by genetically modified mosquitoes may be hindered or aided by dormancy and LDM, as would other forms of malaria control and elimination campaigns.

Divergent strategies of persistence through the dry season were revealed by species-specific seasonal dynamics: local persistence of *A. coluzzii*, as opposed to annual recolonization by LDM for *A. gambiae* s.s. and *A. arabiensis*. They signify a multitude of behavioural, physiological and molecular divergence processes and thus probably represent the most striking phenotypic differences between the species found so far^{28,29}, lending support for the elevated taxonomic status of the molecular forms to species⁶. Consistent with previous interpretations^{6,9,12,19,20}, the adaptation to exploit arid environment such as the Sahel via aestivation may represent the central dimension in the adaptive divergence between the species. The implications of these differences for understanding speciation

and for explaining their geographical range^{11,12} are just beginning to be appreciated.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 3 July; accepted 22 October 2014.

Published online 26 November 2014.

- Donnelly, M. J., Simard, F. & Lehmann, T. Evolutionary studies of malaria vectors. *Trends Parasitol.* **18**, 75–80 (2002).
- Omer, S. M. & Cloudsley-Thompson, J. L. Dry season biology of *Anopheles gambiae* Giles in the Sudan. *Nature* **217**, 879–880 (1968).
- Holstein, M. H. *Biology of Anopheles gambiae. Research in French West Africa* (World Health Organization, 1954).
- Siraj, A. S. et al. Altitudinal changes in malaria incidence in highlands of Ethiopia and Colombia. *Science* **343**, 1154–1158 (2014).
- World Health Organization. *World Malaria Report 2013* (World Health Organization, 2013).
- Coetzee, M. et al. *Anopheles coluzzii* and *Anopheles amharicus*, new members of the *Anopheles gambiae* complex. *Zootaxa* **3619**, 246–274 (2013).
- Omer, S. M. & Cloudsley-Thompson, J. L. Survival of female *Anopheles gambiae* Giles through a 9-month dry season in Sudan. *Bull. World Health Organ.* **42**, 319–330 (1970).
- Taylor, C. E., Toure, Y. T., Coluzzi, M. & Petrarca, V. Effective population size and persistence of *Anopheles arabiensis* during the dry season in west Africa. *Med. Vet. Entomol.* **7**, 351–357 (1993).
- Touré, Y. T. et al. Ecological genetic studies in the chromosomal form Mopti of *Anopheles gambiae* s.s. in Mali, West Africa. *Genetica* **94**, 213–223 (1994).
- Simard, F., Lehmann, T., Lemasson, J. J., Diatta, M. & Fontenille, D. Persistence of *Anopheles arabiensis* during the severe dry season conditions in Senegal: an indirect approach using microsatellite loci. *Insect Mol. Biol.* **9**, 467–479 (2000).
- Coetzee, M., Craig, M. & le Sueur, D. Distribution of African malaria mosquitoes belonging to the *Anopheles gambiae* complex. *Parasitol. Today* **16**, 74–77 (2000).
- della Torre, A., Tu, Z. & Petrarca, V. On the distribution and genetic differentiation of *Anopheles gambiae* s.s. molecular forms. *Insect Biochem. Mol. Biol.* **35**, 755–769 (2005).
- Sogoba, N. et al. Monitoring of larval habitats and mosquito densities in the Sudan savanna of Mali: implications for malaria vector control. *Am. J. Trop. Med. Hyg.* **77**, 82–88 (2007).
- Adamou, A. et al. The contribution of aestivating mosquitoes to the persistence of *Anopheles gambiae* in the Sahel. *Malar. J.* **10**, 151 (2011).
- Huestis, D. L. et al. Seasonal variation in metabolic rate, flight activity and body size of *Anopheles gambiae* in the Sahel. *J. Exp. Biol.* **215**, 2013–2021 (2012).
- Lehmann, T. et al. Aestivation of the African malaria mosquito, *Anopheles gambiae* in the Sahel. *Am. J. Trop. Med. Hyg.* **83**, 601–606 (2010).
- Yaro, A. S. et al. Dry season reproductive depression of *Anopheles gambiae* in the Sahel. *J. Insect Physiol.* **58**, 1050–1059 (2012).
- Lehmann, T. et al. Seasonal variation in spatial distributions of *Anopheles gambiae* in a Sahelian village: evidence for aestivation. *J. Med. Entomol.* **51**, 27–38 (2014).
- Coluzzi, M., Petrarca, V. & Di Deco, M. A. Chromosomal inversion intergradation and incipient speciation in *Anopheles gambiae*. *Boll. Zool.* **52**, 45–63 (1985).
- Coluzzi, M., Sabatini, A., Petrarca, V. & Di Deco, M. A. Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Trans. R. Soc. Trop. Med. Hyg.* **73**, 483–497 (1979).
- Sogoba, N. et al. Spatial distribution of the chromosomal forms of *Anopheles gambiae* in Mali. *Malar. J.* **7**, 205 (2008).
- Toure, Y. T. et al. Perennial transmission of malaria by the *Anopheles gambiae* complex in a north Sudan Savanna area of Mali. *Med. Vet. Entomol.* **10**, 197–199 (1996).
- Denlinger, D. L. Dormancy in tropical insects. *Annu. Rev. Entomol.* **31**, 239–264 (1986).
- Denlinger, D. L. & Armbruster, P. A. Mosquito diapause. *Annu. Rev. Entomol.* **59**, 73–93 (2014).
- Tauber, M. J., Tauber, C. A. & Masaki, S. *Seasonal Adaptations of Insects* (Oxford Univ. Press, 1986).
- Sogoba, N. et al. Malaria transmission dynamics in Niono, Mali: the effect of the irrigation systems. *Acta Trop.* **101**, 232–240 (2007).
- Lemasson, J. J. et al. Comparison of behavior and vector efficiency of *Anopheles gambiae* and *An. arabiensis* (Diptera: Culicidae) in Barkedji, a Sahelian area of Senegal. *J. Med. Entomol.* **34**, 396–403 (1997).
- Huestis, D. L. & Lehmann, T. Ecophysiology of *Anopheles gambiae* s.l.: persistence in the Sahel. *Infect. Genet. Evol.* <http://dx.doi.org/10.1016/j.meegid.2014.05.027> (14 June 2014).
- Lehmann, T. & Diabate, A. The molecular forms of *Anopheles gambiae*: a phenotypic perspective. *Infect. Genet. Evol.* **8**, 737–746 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the residents of Thierola for their hospitality and assistance with mosquito collections; J. Ribeiro, T. Wellems, P. McQueen, R. Faiman and G. Wasserberg for their comments on previous versions of this manuscript; and C. Traoré, R. Sakai, R. Gwadz and T. Wellems for logistical support. This study was supported by the Tamaki Foundation and by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Author Contributions T.L. conceived the study and together with A.D. and A.S.Y. designed it. A.D., A.S.Y., M.D., S.T., D.L.H., Y.K., A.I.T., Z.L.S. and D.S. performed the research, both in the field and the laboratory. All authors have discussed and interpreted the results as well as made decisions on various field and laboratory operations. A.D. led the field operations and data management; T.L. analysed the data and wrote the paper, with extensive input from A.D. and D.L.H.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.L. (tlehmann@niaid.nih.gov).

METHODS

The study was performed between September 2008 and August 2013 in Thierola (13.6583° N, 7.2155° W), a small rural village in the Malian Sahel. The village populations, ethnic composition, agricultural activities and house structure were described previously¹⁸. During the wet season, the rains fill two large ponds and many small puddles near the village. The small puddles require frequent rains, as they dry within a week without additional rain. The last rain typically falls in October and usually all surface water dries by December. From November until May, rainfall is altogether absent or negligible (total precipitation <30 mm). In the course of this study, dry-season 'mango' rains (<20 mm) fell in the area in March 2009, but no rain fell during the dry season in any of the subsequent years, at least over a 30 km radius. During the dry season, water is only available in four deep wells (~25 m deep). Seepage of water around wells and troughs for animals was monitored every dry season, but no mosquito larvae were found in these small puddles, which typically dry up every evening. A few trees may be irrigated by bucket every several days, but all water dries within hours. Annual precipitation is approximately 500 mm (543 mm in Segou, which lies 30 km south and 100 km east of Thierola). For this study, the dry season refers to December–May and the wet season to July–October; the transition periods (June and November) are marked by climatic irregularity (surface water may or may not be available). In this paper, a year is defined as the period spanning from after the first rains (1 July) to the end of the following dry season (30 June). On-the-ground searches for surface waters during the dry season were conducted every year, in consultation with herdsman and hunters, and a detailed examination of the satellite photographs available in Google Earth was also performed. Tree holes containing water that last until January (and rarely into February) were also monitored, but no anopheline larvae were found. Except after the mango rains of March 2009¹⁶, no surface waters have been found during the dry season in a distance up to 30 km around Thierola.

Mosquito collection. Live collections using mouth aspirators inside all houses ($n \approx 120$), were conducted throughout the study period as described previously¹⁸. The number of houses sampled ($n = 511$; median = 119; 95% CI = 103–125) varied because houses were not accessible when their owners were away from the village (and the actual number of houses changed over the five-year study as some were destroyed and others were built). Typically, collections were made every day (dry season) or every-other day (wet season) for two weeks per month. Each house was visited by two trained collectors, both searching for mosquitoes for 10–15 min (and until no mosquitoes were collected for 3–5 min). The same collectors were used throughout the study and rotated across all houses. During certain periods (for example, the wet season of 2008 and 2009, dry season of 2010, and dry season of 2012), collected mosquitoes were marked and released about 1 h after sunset on the day of collection. During other periods, mosquitoes were not released after capture but used for various experiments (reported separately). Because the recapture rate was low (<3%), the effect of removing mosquitoes on the subsequent density, as opposed to releasing them (that is, sampling with and without replacement), was assumed to be negligible.

Additional methods used to collect mosquitoes outdoors included clay pots (with or without water/sugar), CDC light traps (developed by the Centers for Disease Control), fruit/flower baited traps, oviposition traps, emergence traps from larval sites, fence traps, and traps over domestic animals (calves, goats, sheep and chickens), pit latrines, wells or rodent burrows. Although some of these traps were useful during the wet season (for example, emergence traps), they all yielded virtually no mosquitoes during the dry season, as opposed to indoor collections (above), and therefore were discontinued after various intervals (ranging from weeks to years). Several of these methods were described previously^{16,17}. To evaluate congruence between adult and larval composition, larval collections using dippers were conducted during the wet season of 2009 and 2010 from multiple larval sites and multiple positions in each site.

Occasionally other anopheline species were collected by the different methods, including *Anopheles rufipes* and *Anopheles pharoensis*, but their numbers were insufficient for analysis. *A. funestus* was also observed in small numbers during the first year of the study¹⁶ (2008–2009), but virtually vanished during the subsequent years, presumably as a result of the mass distribution of insecticide-impregnated bed nets that started in 2008 in the region.

Data analysis. The *Anopheles gambiae* s.l. indoor collection records, consisting of 511 collection days in all (~120) accessible houses from September 2008 through August 2013, were used to produce mean daily density per house (dividing the total collected (including female and male mosquitoes) in each day by the number of houses searched). Mean daily density per house was transformed to stabilize the variance into natural log density as follows: $\ln(\text{density}) = \text{mean density} + [0.9/(\text{no. of houses sampled})]$. Although statistically equivalent to a transformation using 1, 0.9 was used to signify that the drop of density from 1 to zero, which probably reflects, biologically, 'more' than a change from 1.1 to 0.1.

To provide a descriptive framework as part of the data exploration, putative elements of the annual (seasonal) cycle were visually identified if they appeared in two or more years, after scanning graphs using linear and logarithmic scales (Fig. 1, Extended Data Fig. 2 and Extended Data Table 1) using two 15-day-long frames that were shifted horizontally along the figures. The time series of daily density were converted into an equidistant, 5-day interval time series consisting of 362 intervals using Proc Expand³⁰, after fitting a continuous curve to the data by connecting successive straight-line segments between non-missing input values using the 'join' method. This procedure interpolated missing values in the time series. The 5-day interval estimates produced showed the best fit to the observed data based on visual inspection and the sum of the difference between the observed and expected values, when compared with 10-day, 14-day, and 1 month intervals. Moreover, if the interpolated (missing) value differed from the corresponding 10-day mean by more than 30% of that mean, it was replaced by the latter. If no 10-day mean density was available for that period, the same was carried out with the corresponding global 10-day mean density (across the five years). Less than 5% of values required such substitutions. The fit between the equidistant time series and the observed daily mean density is depicted in Extended Data Fig. 2. Statistical analyses were performed on the equidistant log-density time series.

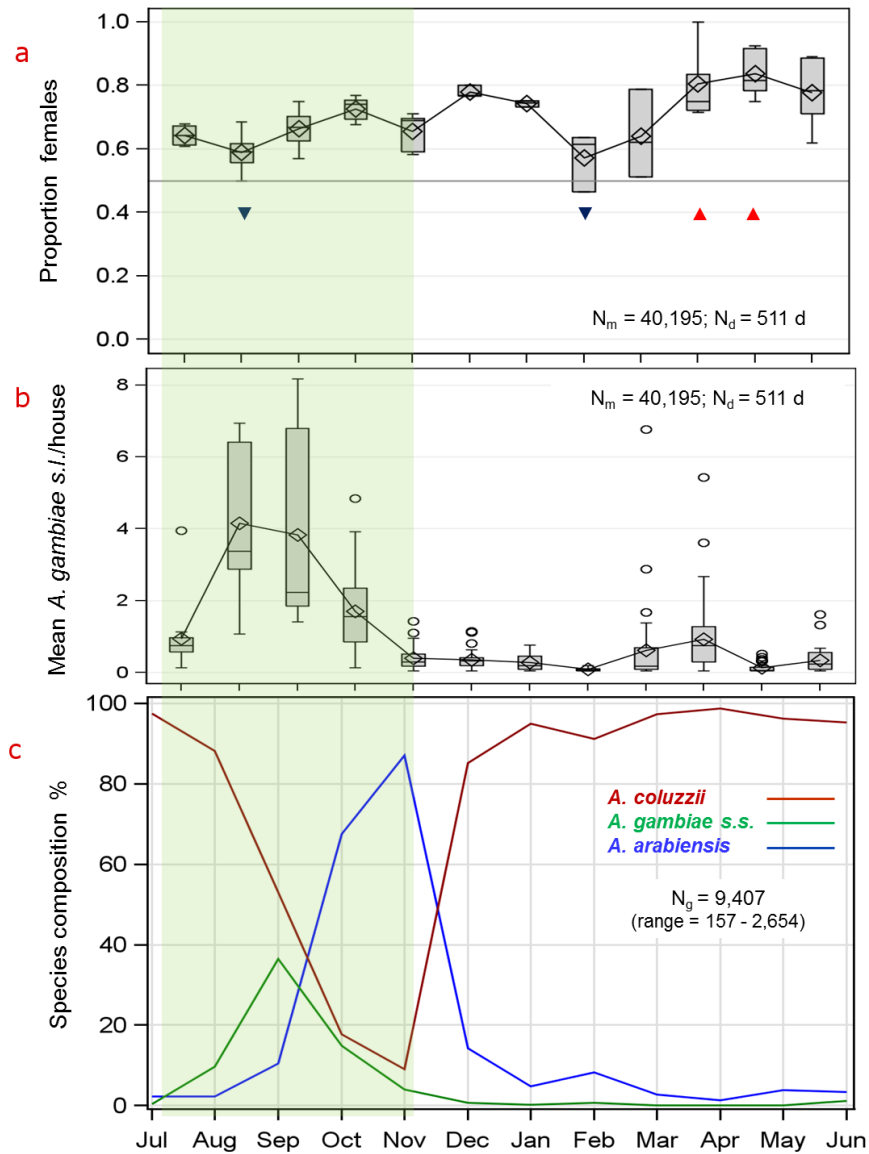
Mosquitoes that were morphologically identified as members of *A. gambiae* s.l. complex were subjected to molecular identification to determine their species³¹. To estimate species composition, we pooled specimens collected from Thierola and nearby villages (up to a 6 km radius) into 10-day intervals based on the day of the month (1–10, 11–20 and 21–31) and separately into monthly intervals. In a few cases with small sample sizes ($n < 15$), we pooled two consecutive 10-day intervals. The resulting series had variable gaps representing missing values either because no mosquitoes were collected despite extensive collection effort (for example, January–February 2012), or because no collections were made (for example, December 2008–March 2009). When no composition data were available for the whole month, the monthly mean fractions of each species estimated for that month across the five years were imputed for the missing values. Composition data for 10/61 months (16%) were imputed in that way. The compositional series consisting of 10 d estimates and imputed monthly values were thereafter interpolated using Proc Expand³⁰ to 10-day intervals (without changing observed compositional values) for each species separately. The interpolated values were restricted to values between 0 and 100. The species-specific (absolute) mean density was then estimated as the product of the proportion of each species at that 10-day time interval by the density of *A. gambiae* s.l. at the corresponding 5-day time intervals described above.

Time-series analysis of the log density of each species was carried out using the Unobserved Components Model in SAS³⁰ (Proc UCM), which accommodates time-varying parameters of the trend, seasonal, and cycle components derived from decomposition of the time series, as well as various methods of incorporating autoregressive processes. It estimates both deterministic and stochastic parameters and provides tests of the parameters' variance to determine if these parameters are time-varying. Overall goodness-of-fit measures, such as Akaike information criterion (AIC) used to compare models were computed, as well as extensive tests of the residuals and diagnostic graphics. We tested whether overall seasonal variation was statistically significant and if so, determined if it was time-varying or constant, before identifying its salient elements. The seasonal component is a unique cycle with a strictly annual periodicity (whose parameters sum to zero over a year). Seasonality was modelled as a series of 73 dummy variables, each representing a 5-day interval. Starting with the basic structural model³² that includes slope, level, and seasonal components (all stochastic, in addition to the irregular element), we removed or added one parameter at a time and evaluated the significance of all parameters, the overall fit of the model, and the residual diagnostics for serial correlation, heteroscedasticity, and normality. Additional cycles or autoregressive functions may be required to model the interdependencies of the data between time-points until the distribution of the residuals complies with white noise (Supplementary Information). This approach led to selecting a parsimonious model that accounted well for the pattern of the time series and met the required assumptions. The seasonal component extracted from the selected model and its 95% CI were used to identify elements (phases) with statistical support. Thus, a peak whose 95% CI did not overlap with its adjacent minima had statistical support. All tests and *P*-values are based on two-sided tests.

30. SAS (Sas Institute, Cary, North Carolina, 2011).

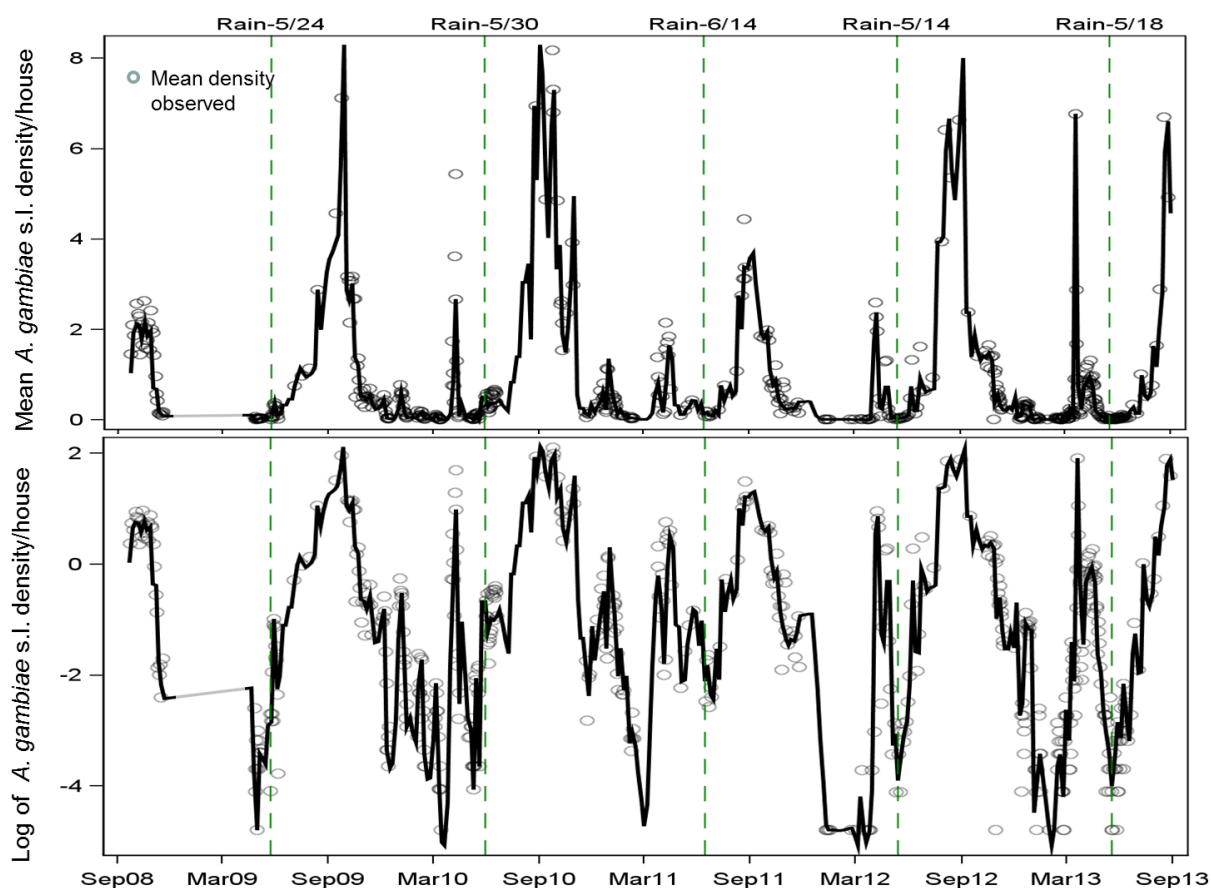
31. Fanelli, C., Santolamazza, F. & della Torre, A. Simultaneous identification of species and molecular forms of the *Anopheles gambiae* complex by PCR-RFLP. *Med. Vet. Entomol.* **16**, 461–464 (2002).

32. Harvey, A. C. *Forecasting, Structural Time Series Models and the Kalman Filter* (Cambridge Univ. Press, 1989).



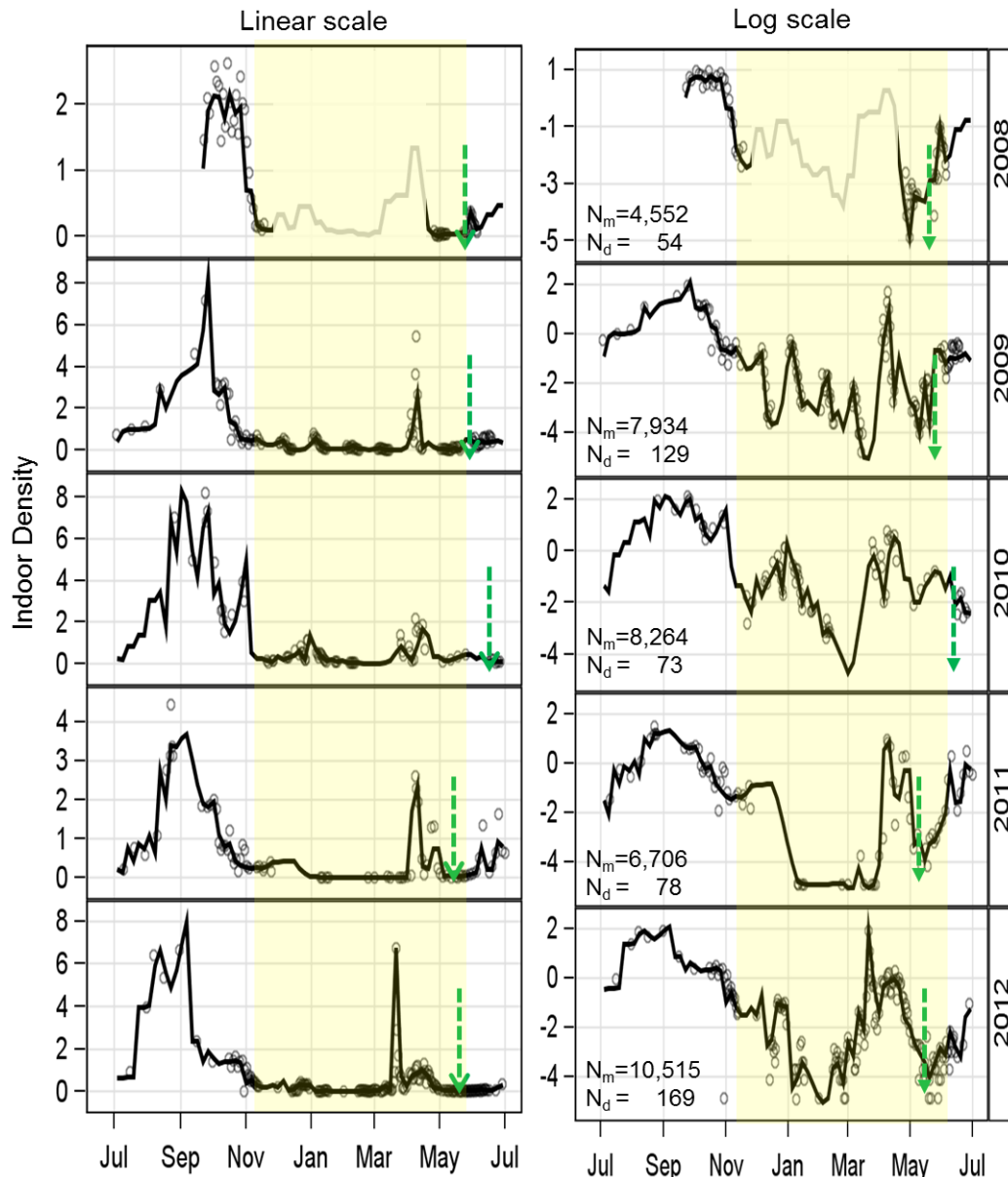
Extended Data Figure 1 | Sex-ratio, density and composition of *Anopheles gambiae* s.l. **a–c**, Overall monthly means of the proportion of *A. gambiae* s.l. females (**a**), house density (**b**), and species composition (**c**). N_m , N_d , and N_g denote sample size of *A. gambiae* s.l., the number of collection days, and the number genotyped to species, respectively (Methods). Whiskers in

box-whisker plots extend to the extreme values up to 1.5 the distance between the twenty-fifth and seventy-fifth percentiles. In **a**, blue triangles represent means that are significantly lower than the red triangles (based on the sequential Bonferroni test; see Extended Data Table 1) and the horizontal line represents 1:1 sex ratio.



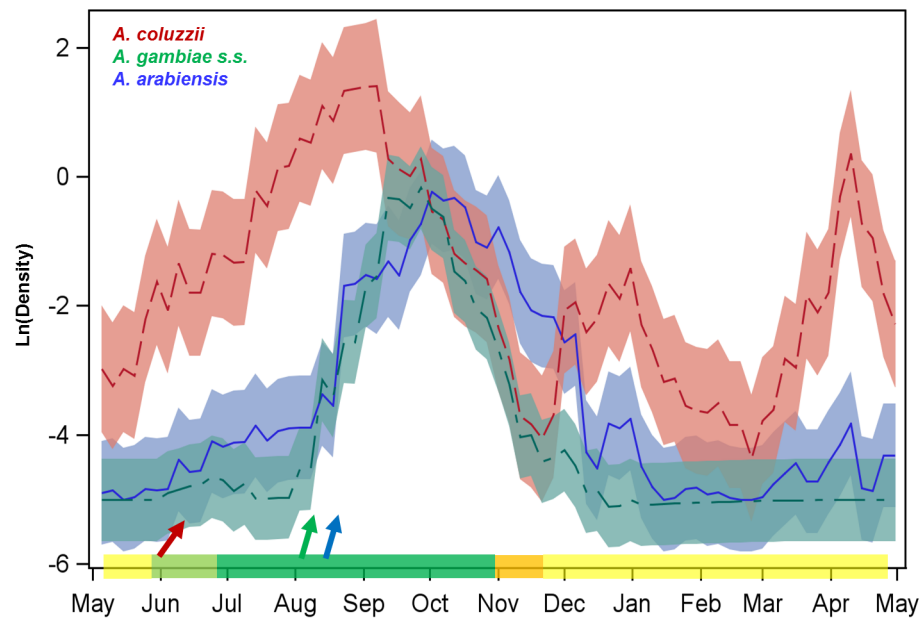
Extended Data Figure 2 | Population dynamics of *Anopheles gambiae s.l.* House density over time in linear (top) and natural logarithm (bottom) scale to evaluate systematic change over time. Circles denote observed daily mean density and the black lines show the interpolated series of mean house density

over 5-day intervals. Grey lines depict interpolation during the longest time without field samples (December 2008 to April 2009). First rain events are shown by green lines (dates listed above). Sample sizes of mosquitoes and collection days are the same as in Extended Data Fig. 1b.



Extended Data Figure 3 | Population dynamics of *Anopheles gambiae* s.l. across years. Observed daily mean density (circles) is shown against 5-day means (line) on linear and log scales from July to June of every year to assess similarity among years. Green arrows mark the first rain and the tan

background denotes the dry season. Shading during 2008 indicates a gap in sampling (December–March) when imputed values were used (Methods). Sample sizes (N_m and N_d) are explained in Fig. 1.



Extended Data Figure 4 | Level-adjusted seasonal component of the density of *A. coluzzii*, *A. gambiae s.s.* and *A. arabiensis*. The level, seasonals, and their 95% CI (bands) were estimated using unobserved component time series models (Table 1; Methods). Arrows denote the start of the population

growth (text and Fig. 2). Time is shown starting from May to maximize the comparability between the species. In the colour ruler on the x axis, yellow and green denote dry and wet seasons, respectively, and orange and light-green denote transition periods. Sample sizes are given in Extended Data Fig. 1.

Extended Data Table 1 | Annual and monthly variation in the indoor sex ratio (proportion of females)

Source	df	F/Z	P
Year	--	0.34	0.37
Month	11/33	5.35	0.0001
AR(1)	--	0.39	0.69
Residual	--	4.12	0.0001
AIC/-2ResLL		-65/59	

Month was treated as fixed effect and year as random effect in Proc Mixed³⁰. No indication for serial correlation was detected.

Extended Data Table 2 | Putative elements of the seasonal cycles of the members of *A. gambiae* s.l.

Taxon ^a	Year	Wet season (WS)			Dry season (DS)								WS
		normal growth	WS peak	late-WS low	rapid decline	slow decline		mid DS low	DS rise	late-DS peak	late-DS decline	end-DS low	first-rain surge
		Jun-Aug	Sep-Oct	Oct	Nov	Dec-Feb		Feb-Mar	Mar	Apr	Apr	May	May-Jun
<i>A. gambiae</i> s.l.	2008-09	nd	Y	Y	Y	nd		nd	nd	nd	nd	Y	Y
	2009-10	Y	Y	Y	Y	Y		Y	nd	Y	N	Y	Y
	2010-11	nd	Y	Y	Y	Y		Y	nd	Y	Y	Y	N
	2011-12	Y	Y	Y	Y	N		Y	Y	Y	Y	Y	Y
	2012-13	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y	N
Taxon ^a	Year	normal growth	WS peak	WS decline	pre DS trough	early-DS peak	early-DS decline	mid-DS low	DS rise	late-DS peak	late-DS decline	end-DS low	first-rain surge
		Jun-Jul	Aug-Sep	Sep-Nov	Nov-Dec	Dec	Jan-Feb	Feb-Mar	Mar	Apr	Apr	May	May-Jun
<i>A. coluzzii</i>	2008-09	nd	Y	Y	Y	nd	nd	nd	nd	nd	nd	Y	Y
	2009-10	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
	2010-11	nd	Y	Y	Y	Y	Y	Y	nd	Y	Y	Y	N
	2011-12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	2012-13	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Taxon ^a	Year	normal growth	WS peak	decline		DS desertion							onset
		Aug	Sep-Oct	Nov-Dec		Jan-May							Jun-Jul
<i>A. gambiae</i> s.s.	2008-09	nd	Y	Y		nd							Y
	2009-10	Y	Y	Y		Y							Y
	2010-11	nd	Y	Y		Y							N
	2011-12	Y	Y	Y		Y							N
	2012-13	Y	Y	Y		Y							Nd
Taxon ^a	Year	normal growth	WS peak	decline		DS sporadic appearance							
		Aug-Sep	Oct	Nov-Dec		Jan-Jul							
<i>A. arabiensis</i>	2008-09	nd	Y	Y		nd							
	2009-10	Y	Y	Y		Y							
	2010-11	nd	Y	Y		Y							
	2011-12	Y	Y	Y		N							
	2012-13	Y	Y	Y		Y							

Elements were identified based on visual examination of their population dynamics over the five years of the study (Methods, Fig. 1 and Extended Data Fig. 3). 'Y', 'N', and 'nd', denote periods when a putative element was visible, invisible, or was not determined (because of insufficient data), respectively.

^a Each taxon has its own 'element header' with the typical month in which an element was observed. Species-specific elements are shown in colour. Underlined elements were supported statistically (text, Methods and Fig. 2).

A relative shift in cloacal location repositions external genitalia in amniote evolution

Patrick Tschopp¹, Emma Sherratt^{2†}, Thomas J. Sanger^{2†}, Anna C. Groner³, Ariel C. Aspiras¹, Jimmy K. Hu^{1†}, Olivier Pourquie^{1,4,5}, Jérôme Gros⁶ & Clifford J. Tabin¹

The move of vertebrates to a terrestrial lifestyle required major adaptations in their locomotory apparatus and reproductive organs. While the fin-to-limb transition has received considerable attention^{1,2}, little is known about the developmental and evolutionary origins of external genitalia. Similarities in gene expression have been interpreted as a potential evolutionary link between the limb and genitals^{3–6}; however, no underlying developmental mechanism has been identified. We re-examined this question using micro-computed tomography, lineage tracing in three amniote clades, and RNA-sequencing-based transcriptional profiling. Here we show that the developmental origin of external genitalia has shifted through evolution, and in some taxa limbs and genitals share a common primordium. In squamates, the genitalia develop directly from the budding hindlimbs, or the remnants thereof, whereas in mice the genital tubercle originates from the ventral and tail bud mesenchyme. The recruitment of different cell populations for genital outgrowth follows a change in the relative position of the cloaca, the genitalia organizing centre. Ectopic grafting of the cloaca demonstrates the conserved ability of different mesenchymal cells to respond to these genitalia-inducing signals. Our results support a limb-like developmental origin of external genitalia as the ancestral condition. Moreover, they suggest that a change in the relative position of the cloacal signalling centre during evolution has led to an altered developmental route for external genitalia in mammals, while preserving parts of the ancestral limb molecular circuitry owing to a common evolutionary origin.

To investigate potential interdependencies between the development of the limbs and external genitalia, we first determined the location of the two structures during embryogenesis. We focused on mouse^{4,5} and squamates (lizard and snakes), which show progressive limb reduction⁷ yet maintain their external genitalia, the hemipenes⁸. Micro-computed-tomography (μ CT) reconstructions of mouse, anole lizard (*Anolis*), python and house snake embryos revealed different anterior–posterior locations of the developing external genitalia relative to limbs. In mice, the genital tubercle is positioned caudal to the hindlimbs (Fig. 1a), whereas in squamates the paired hemipenes bud from the limbs, or from the remnants thereof (Fig. 1b–d). The cloaca, a signalling centre important for genitalia development^{6,9,10}, is similarly located within the limb-field of squamates (Fig. 1f–h) and expresses *Shh* (Fig. 1j–l). Thus, in squamates all three anatomical structures—limb, hemipenis and cloaca—align at the same anterior–posterior position.

We decided to investigate whether these positional differences would reflect distinct developmental origins of external genitalia. Although the cells of the vertebrate limb bud are known to arise through an epithelial-to-mesenchymal transition (EMT) of an epithelial lateral plate mesoderm (LPM) population lining the coelomic cavity¹¹, the developmental origin of external genitalia in vertebrates is still unclear. We developed a lentiviral lineage tracing approach (see Methods), to systematically follow the two sources previously proposed, the LPM and tail bud^{10,12,13},

in three amniote species: mouse, chicken and anole. Injections into the coelom of embryonic day (E)9.5 mouse embryos label cells surrounding the coelom as well as the developing hindlimb (Fig. 2a). However, no green fluorescent protein (GFP)-positive cells are observed in the genital tubercle, with a sharp boundary of labelled cells extending from the coelomic cavity (Fig. 2a, b). In contrast, injection into the posterior mesenchyme of mouse embryos labels the genital tubercle (Fig. 2c). Tail bud injections label the posterior half, whereas the infra-umbilical mesenchyme gives rise to its anterior part^{12–14} (Extended Data Fig. 1). In chicken, coelomic injection into stage HH14 embryos labels cells in both limb and genital tubercle (Fig. 2d, e), without any obvious boundary. Tail bud infection also results in GFP-positive genitalia cells, mostly in the posterior tubercle, suggesting that multiple lineages contribute to this species' genitalia (Fig. 2f). Similar conclusions were reached in a parallel study³¹. In *Anolis*, coelomic injections at stage 2–3 result in GFP-positive cells in the limb and the developing genitalia (Fig. 2g, h), whereas no labelled cells are seen in the hemipenes after tail bud injections (Fig. 2i).

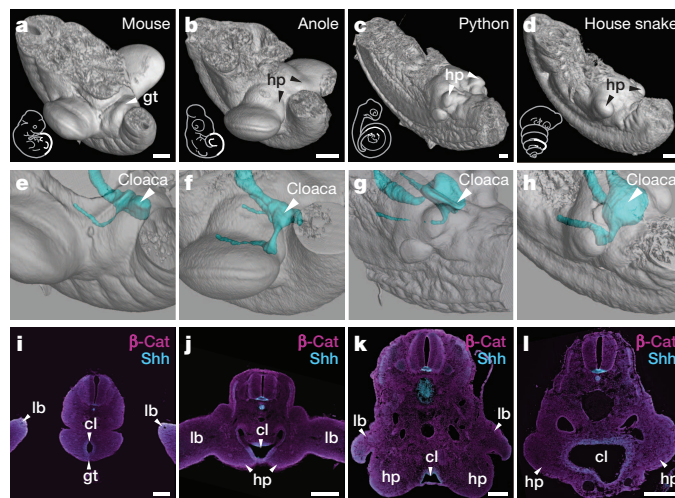


Figure 1 | A relative positional shift of limbs, genitalia and the cloaca in squamates. a–d, μ CT scans of mouse (a), anole (b), python (c) and house snake (d) lumbosacral regions (highlighted in sketch in white) at embryonic stages, illustrating the position of the developing external genitalia. e–h, Three-dimensional reconstructions of cloacal volumes. The cloaca is located at the same anterior–posterior position as the limb in squamates (f–h); however, it is positioned more posteriorly in the mouse (e). i–l, Transversal sections stained for β -catenin and *Shh*, indicating the conservation of a cloacal signalling centre in all four species. gt, genital tubercle; hp, hemipenis; lb, limb; cl, cloaca. Scale bars, 200 μ m.

¹Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. ²Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

³Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. ⁴Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), 67400 Illkirch, France.

⁵Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. ⁶Developmental and Stem Cell Biology Department, Institut Pasteur, 75724 Paris Cedex 15, France.

[†]Present addresses: Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, Iowa 50011, USA (E.S.); Department of Molecular Genetics and Microbiology, University of Florida, Gainesville, Florida 32610, USA (T.J.S.); Department of Orofacial Sciences and Program in Craniofacial and Mesenchymal Biology, UCSF, San Francisco, California 94143, USA (J.K.H.).

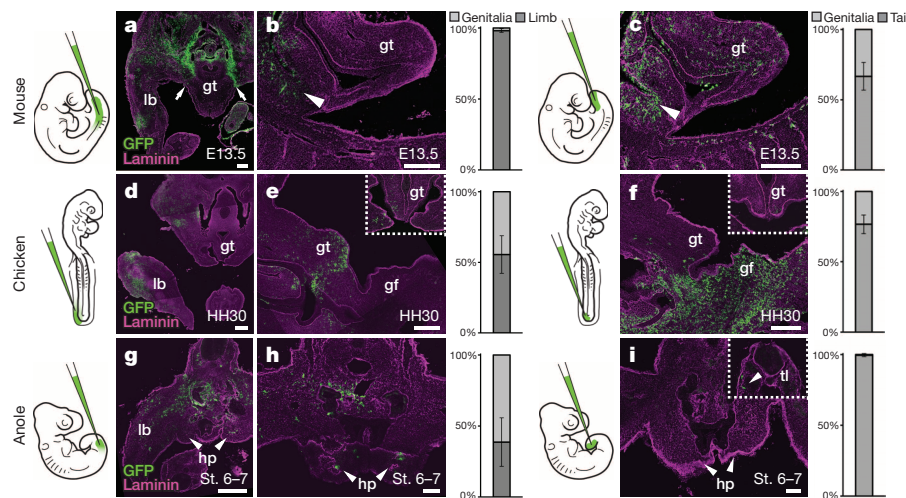


Figure 2 | Differential developmental origins of external genitalia in amniotes. **a–i**, Transversal and sagittal views of GFP lentivirus-injected embryos. Relative contribution of GFP-positive cells to respective organs is quantified on the right, normalized on tissue area. Error bars represent standard deviation in at least $n = 4$ biological replicates. **a, b**, Injection into the coelom at mouse E9.5 ($n = 48$) labels the limb at E13.5, but excludes the genital tubercle (arrows). Only cells lining the peritoneal cavity are labelled (**b**, arrowhead), but none in the genital tubercle proper. **c**, Injection into the tail bud ($n = 101$) labels cells in the genital tubercle. Accidental piercing of the coelom labels cells of the peritoneal cavity (arrowhead). **d, e**, Coelom injection

in HH14 chicken embryos ($n = 81$) labels the limb and the genital tubercle at HH30. **e**, Sagittal and transversal close-up (inset) views. **f**, Sagittal and transversal close-up (inset) views of tail-bud-injected chick embryos ($n = 77$), showing labelling in the genital tubercle. **g, h**, Anole embryos injected into the coelom at stage (St.) 2–3 ($n = 94$) show GFP labelling of both limb and hemipenis at stage 6–7. **i**, No hemipenis cells are labelled following tail bud injection ($n = 57$), even though there are GFP-positive cells in the tail (inset, arrowhead). gt, genital tubercle; gf, genital fold; hp, hemipenis; lb, limb; tl, tail. Scale bars, 200 μm (**a–g**), 50 μm (**h, i**).

As for chicken and mouse limbs¹¹, cells of the *Anolis* limb, but also the hemipenis, originate through an EMT of the coelomic epithelium (Extended Data Fig. 2a, b). In snakes, we find evidence for similar cellular dynamics. The hemipenes emerge as small buds at a ‘limb-like’ lateral position, juxtaposed to the coelomic cavity (Extended Data Fig. 2c–f). A concomitant basement membrane breakdown, consistent with an EMT of the LPM, is seen in the budding of both mouse limbs and snake hemipenes (Extended Data Fig. 2d, f). Moreover, we find that *Tbx4*, a gene important for hindlimb development¹⁵ is expressed from early on, in both the coelomic epithelium and the mesenchyme of the developing hemipenis (Extended Data Fig. 2g–i). Its forelimb counterpart *Tbx5* is expressed later, in the mesenchyme only, in agreement with its pattern of expression in the genitalia of mammals¹⁶ (Extended Data Fig. 2j–l). This suggests that squamate external genitalia initiate with limb-like cellular dynamics, with the resulting mesenchymal cell population in modern snakes being converted to a genital fate¹⁷. Thus, important differences exist in the developmental origins of external genitalia in amniotes: chicken genitals originate from both LPM and the tail bud, whereas the mouse genital tubercle consists of infra-umbilical mesenchyme and tail-bud-descendant cells. In contrast, the hemipenis shares a developmental route with hindlimbs, either through secondary budding as in lizards, or by entirely recruiting the mesenchymal cell population to a genital fate in modern snakes. Given the impact of developmental lineage on an organ’s molecular architecture, we next explored the transcriptomes of emerging genitalia, in the two opposing trajectories of mammals and squamates.

Gene expression profiling has successfully been applied to questions of developmental and evolutionary origin, of cell types and entire morphological features^{18–21}. We thus dissected the early and late stages of developing limbs and genitalia from mouse and anole embryos (Fig. 3a) for comparative RNA-sequencing (RNA-seq) analyses. Overall transcriptome similarities were assessed using multidimensional scaling (MDS; Fig. 3b). The transcriptomes dominantly resolve along dimension 1 in a species-dependent manner, as expected for similar tissues in evolutionarily distant species²². Dimension 2, however, contains a clear organ identity signal, that is, limb versus genitalia. This separation is virtually absent in *Anolis* samples compared with mouse, reflecting similar

transcriptional programs owing to a common developmental origin for genitalia and hindlimbs. For hierarchical clustering, we included the early tail bud as outgroup of the primary body axis, and the forelimbs to account for anterior–posterior differences in the two species’ genitalia (Fig. 3c, d; see Methods). In *Anolis*, the early hemipenis transcriptome falls within the limb clade, indicative of an almost generic limb molecular architecture, and only later differentiates into a more organ-specific signature (Fig. 3c and Extended Data Fig. 3a). In contrast, mouse genitalia transcriptomes are clearly distinct from limbs, from early on, highlighting the separate developmental origins of the two organs (Fig. 3d and Extended Data Fig. 3b).

To identify genes driving hindlimb- and genitalia-specific transcriptome separation, we used principal component analysis. Principal component 1 (PC1) correlates with species differences, whereas organ specificity is resolved along PC2 (Fig. 3e), allowing us to identify organ-specific ‘driver’ genes in a species-independent manner. We assessed the contribution of orthologous genes to PC2, according to their absolute loading values (Fig. 3f). Gene ontology (GO) analysis²³ of the top 500 genes (Fig. 3f and Supplementary Table 1) identified GO terms related to transcription factors and signalling molecules (Fig. 3g). Gene regulatory networks thus determine limb versus genital organ transcriptomes, but also mirror their developmental origin. Notably, transcription factor/signalling molecule data are sufficient to reproduce the clustering seen with whole transcriptomes (Fig. 3h, i and Extended Data Fig. 4a, b). Collectively, we find a clear distinction between mouse genitalia and limb transcriptomes during early and late organogenesis. Such genitalia-specific separation is only seen in late *Anolis* hemipenes, arguing for a developmental repurposing to a copulatory structure. Importantly, we find transcriptional similarities between early hemipenes and hindlimbs in squamates, illustrating a common developmental origin.

An attractive model for the varied developmental origins of amniote external genitalia would be a repositioning of the cloacal signalling centre with respect to different mesenchymal cell populations with progenitor potential. Hence, bringing either hindlimb or tail bud close to the cloaca would allow these lineages to contribute to genital outgrowth. We tested this hypothesis by grafting GFP-transgenic chicken or quail cloacae into the hindlimb bud of wild-type embryos (Fig. 4a and Extended Data

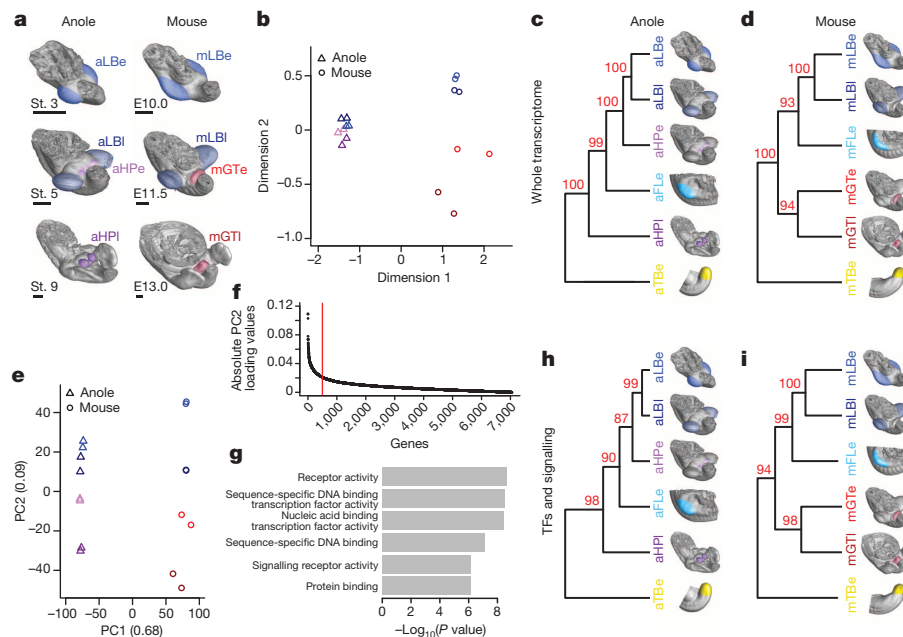


Figure 3 | Molecular architecture of limbs and external genitalia in lizards and mice. **a**, Micro-dissected tissues for RNA-seq analysis, highlighted by colour code and four-letter sample identifier. Early and late limb buds and genitalia buds were analysed in anole lizard and mouse embryos ($n = 2$). **b**, MDS analysis reveals greater overall transcriptome similarities in anole limb and genitalia data sets (triangles) than in their mouse counterparts (circles). **c**, **d**, Hierarchical clustering of pairwise Pearson's correlation coefficients for whole-transcriptome data from anole (**c**) and mouse (**d**) samples. Additional data sets are stage (St.) 2–3 anole and E9.5 mouse forelimb (turquoise) and tail bud (yellow). Numbers at nodes represent approximately unbiased P values,

obtained by multiscale bootstrap resampling. **e**, Principal component analysis. Species transcriptomes separate along PC1, whereas organs are resolved along PC2. Numbers in brackets indicate proportion of variance, explained by PC1 and PC2. **f**, Absolute loading values for PC2, as shown in **e**. **g**, GO term enrichment analysis using the top 500 genes (red line in **f**). Top hits include transcription-factor- and signalling-pathway-related terms. **h**, **i**, Hierarchical clustering analysis of pairwise Pearson's correlation coefficients for transcription factors (TFs) and signalling pathways data from anole (**h**) and mouse (**i**) samples. Sample identifiers: a, anole; m, mouse; GT, genital tubercle; HP, hemipenis; LB, limb; e, early; l, late.

Fig. 5a–c). After 1–2 days of incubation, limbs showed ectopic, secondary buds (Fig. 4b, c). GFP-negative cells suggest an inductive effect of the cloaca on the surrounding mesenchyme, rather than simple over-proliferation of the graft itself (Fig. 4c). Similar buds could be induced by grafting beads soaked in the known cloacal signalling molecules

SHH and FGF^{9,10,24,25} (Extended Data Fig. 5d–g). To assess the fate of responding cells, we re-analysed our RNA-seq data for potential genital versus limb markers. For both species, we performed stage- and organ-specific differential expression analyses. Of the 2,003 genes showing an absolute log₂ fold change greater than 1.5 (P value < 0.05) (Extended Data Fig. 6), we identified 27 that are altered in all four comparisons, 25 of which were altered in the same direction (Fig. 4d). Hierarchical clustering of normalized mouse expression values reveals four stage- and organ-specific signatures, which are largely conserved in *Anolis* (Fig. 4e). On the basis of expression patterns and levels (Extended Data Figs 6 and 7), we chose marker genes to assess transcriptional changes due to ectopic cloacal signals. Indeed, chicken limb cells close to GFP-positive cloacal grafts downregulate limb markers *LHX9* and *TBX18* (Fig. 4f, g), and ectopically express genital markers *ISL1* (ref. 26), *GATA2* and *RUNX1*

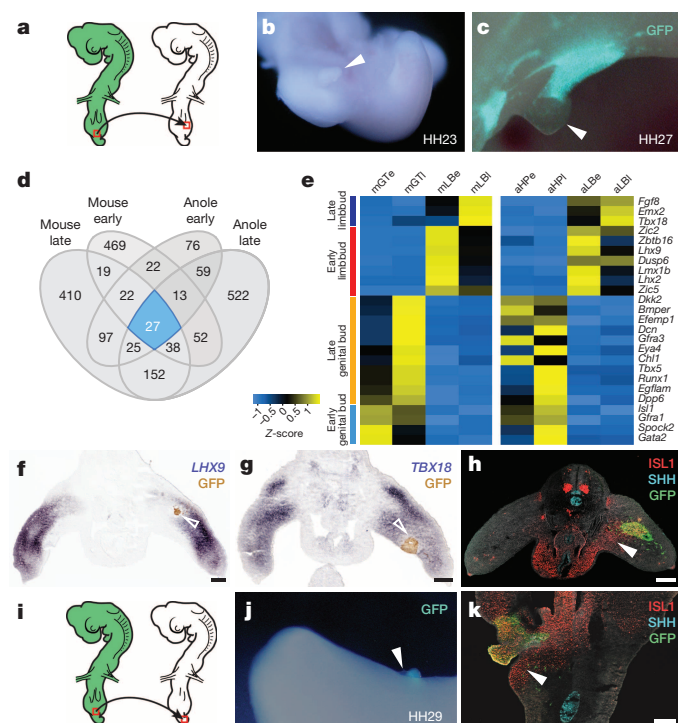


Figure 4 | The cloacal signalling centre can recruit different mesenchymal cell populations for the outgrowth of external genitalia. **a**, Schematic of the hindlimb grafting procedure in chicken embryos. GFP-transgenic cloacae are transplanted into the proximal-ventral portion of wild-type hindlimbs. **b**, **c**, Ectopic outgrowth (arrowheads) on limbs with cloacal grafts ($n = 30/118$). **d**, Venn diagram of pairwise differential expression analysis results (log(fold change) > 1.5, P value < 0.05) of limbs versus genital tissues for early and late budding stages in anole and mouse. **e**, Heat map of Z-score-normalized values of core 25 genes showing consistent differential expression between limbs and genitalia. Mouse row-based hierarchical clustering was re-used for anole samples. **f–h**, Analysis of limb- and genital-specific markers. Expression of limb markers *LHX9* (**f**) and *TBX18* (**g**) is downregulated near the GFP-positive cloacal graft (open arrowheads), while genital marker *ISL1* (**h**) is expressed ectopically (arrowhead). **i**, Schematic of the tail-bud-grafting procedure in chicken embryos. GFP-transgenic cloacae are transplanted into ventral wild-type tail buds. **j**, Ectopic outgrowth on tails with cloacal grafts ($n = 16/87$). **k**, Genital marker *ISL1* is upregulated ectopically in the tail, close to the GFP-positive cloacal graft (arrowhead). All gene expression was assessed in at least $n = 3$ samples. Scale bars, 200 μ m.

(Fig. 4h and Extended Data Fig. 5h–k). Accordingly, when we graft cloacal tissue into the tail bud mesenchyme, ectopic budding and genital marker expression are equally induced (Fig. 4i–k and Extended Data Figs 5l–n and 8). Hence, mesenchymal cells from different developmental origins can respond to inductive cloacal signals, generating outgrowths and genitalia-like marker gene expression. Importantly, these results support the idea that changing the relative anterior–posterior position of the cloaca could generate external genitalia with distinct developmental origins during the course of amniote evolution.

In summary, we show substantial variation in external genitalia development in extant amniote species. We propose that repositioning the cloaca can recruit different mesenchymal cell populations, either through spatial or heterochronic changes in the dynamics of their emergence. In squamates, the hindlimb is the dominant source, with modern snakes entirely repurposing a mesenchymal bud to a genital fate. In mice, limbs and genitals have discrete developmental origins—the LPM and the ventral and tail bud mesenchyme, respectively—with chicken showing an intermediate state. Moreover, we find that similarities in limb and genitalia transcriptomes are dependent on the cellular source of the primordia from which they emerge. Specifically, there is a higher degree of early transcriptome congruence in species deriving their intromittent organs from limb anlagen. Notably, the ability of different mesenchymal cell populations to respond to cloacal, genitalia-inducing signals seems conserved in extant species. It is therefore tempting to speculate that a limb-derived state could represent the ancestral condition in the evolution of external genitalia, as suggested by their position relative to limbs during turtle development^{17,27} and the bifid genitalia of basal mammals^{28,29}. As such, a developmental continuity between limbs and genitalia could have turned into an ‘evolutionary continuity’ in mammals, as the two organs spatially separated owing to a relative repositioning of the cloaca³⁰. Once-shared developmental trajectories could thus help to explain molecular similarities still noticeable in species that now develop the limb and genitalia from distinct cellular sources^{3,4,19}.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 March; accepted 20 August 2014.

Published online 5 November; corrected online 17 December 2014 (see full-text HTML version for details).

- Shubin, N., Tabin, C. & Carroll, S. Fossils, genes and the evolution of animal limbs. *Nature* **388**, 639–648 (1997).
- Wagner, G. P. & Chiu, C. H. The tetrapod limb: a hypothesis on its origin. *J. Exp. Zool.* **291**, 226–240 (2001).
- Kondo, T., Zakany, J., Innis, J. W. & Duboule, D. Of fingers, toes and penises. *Nature* **390**, 29 (1997).
- Yamada, G. *et al.* Molecular genetic cascades for external genitalia formation: an emerging organogenesis program. *Dev. Dyn.* **235**, 1738–1752 (2006).
- Cohn, M. J. Development of the external genitalia: conserved and divergent mechanisms of appendage patterning. *Dev. Dyn.* **240**, 1108–1115 (2011).
- Lin, C. *et al.* Delineating a conserved genetic cassette promoting outgrowth of body appendages. *PLoS Genet.* **9**, e1003231 (2013).
- Greer, A. E. Limb reduction in squamates: identification of the lineages and discussion of the trends. *J. Herpetol.* **25**, 166–173 (1991).
- Cope, E. D. On the hemipenes of the Sauria. *Proc. Acad. Nat. Sci. Philadelphia* **48**, 461–467 (1896).
- Haraguchi, R. *et al.* Unique functions of sonic hedgehog signaling during external genitalia development. *Development* **128**, 4241–4250 (2001).
- Perriton, C. L., Powles, N., Chiang, C., Maconochie, M. K. & Cohn, M. J. Sonic hedgehog signaling from the urethral epithelium controls external genital development. *Dev. Biol.* **247**, 26–46 (2002).
- Gros, J. & Tabin, C. J. Vertebrate limb bud formation is initiated by localized epithelial-to-mesenchymal transition. *Science* **343**, 1253–1256 (2014).

- Ohta, S., Suzuki, K., Tachibana, K., Tanaka, H. & Yamada, G. Cessation of gastrulation is mediated by suppression of epithelial-mesenchymal transition at the ventral ectodermal ridge. *Development* **134**, 4315–4324 (2007).
- Suzuki, K., Economides, A., Yanagita, M., Graf, D. & Yamada, G. New horizons at the caudal embryos: coordinated urogenital/reproductive organ formation by growth factor signaling. *Curr. Opin. Genet. Dev.* **19**, 491–496 (2009).
- Matsumaru, D. *et al.* Genetic analysis of the role of *Alx4* in the coordination of lower body and external genitalia formation. *Eur. J. Hum. Genet.* **22**, 350–357 (2014).
- Naiche, L. A. Loss of *Tbx4* blocks hindlimb development and affects vascularization and fusion of the allantois. *Development* **130**, 2681–2693 (2003).
- Chapman, D. L., Garvey, N., Hancock, S. & Alexiou, M. Expression of the T-box family genes, *Tbx1–Tbx5*, during early mouse development. *Dev. Dyn.* **206**, 379–390 (1996).
- Raynaud, A. & Pieau, C. in *Biology of the Reptilia, Development B* (eds Gans, C. & Billett, F. S.) Vol. 15, 149–300 (John Wiley & Sons, 1985).
- Arendt, D. Genes and homology in nervous system evolution: comparing gene functions, expression patterns, and cell type molecular fingerprints. *Theory Biosci.* **124**, 185–197 (2005).
- Wagner, G. P. The developmental genetics of homology. *Nature Rev. Genet.* **8**, 473–479 (2007).
- Shubin, N., Tabin, C. & Carroll, S. Deep homology and the origins of evolutionary novelty. *Nature* **457**, 818–823 (2009).
- Wang, Z., Young, R. L., Xue, H. & Wagner, G. P. Transcriptomic analysis of avian digits reveals conserved and derived digit identities in birds. *Nature* **477**, 583–586 (2011).
- Merkin, J., Russell, C., Chen, P. & Burge, C. B. Evolutionary dynamics of gene and isoform regulation in mammalian tissues. *Science* **338**, 1593–1599 (2012).
- Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. *Nature Genet.* **25**, 25–29 (2000).
- Haraguchi, R. *et al.* Molecular analysis of external genitalia formation: the role of fibroblast growth factor (*Fgf*) genes during genital tubercle formation. *Development* **127**, 2471–2479 (2000).
- Miyagawa, S. *et al.* Dosage-dependent hedgehog signals integrated with Wnt/ β -catenin signaling regulate external genitalia formation as an appendicular program. *Development* **136**, 3969–3978 (2009).
- Suzuki, K. *et al.* Reduced BMP signaling results in hindlimb fusion with lethal pelvic/urogenital organ aplasia: a new mouse model of sirenomelia. *PLoS ONE* **7**, e43453 (2012).
- Larkins, C. E. & Cohn, M. J. Phallus development in the turtle *Trachemys scripta*. *Sex Dev.* <http://dx.doi.org/10.1159/000363631> (2014).
- Griffiths, M. in *The Biology of the Monotremes* (Academic, 1978).
- De Barros, M. A. *et al.* Marsupial morphology of reproduction: South America opossum male model. *Microsc. Res. Tech.* **76**, 388–397 (2013).
- Jurberg, A. D., Aires, R., Varela-Lasheras, I., Novoa, A. & Mallo, M. Switching axial progenitors from producing trunk to tail tissues in vertebrate embryos. *Dev. Cell* **25**, 451–462 (2013).
- Herrera, A. M. & Cohn, M. J. Embryonic origin and compartmental organization of the external genitalia. *Sci. Rep.* **4**, 6896 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements The authors thank D. Duboule, H. Kaessmann, A. Necseulea, B. Okaty, G. Rey and G. P. Wagner for discussions, M. A. de Bakker for the snake *Tbx5* probe and A. M. Herrera and M. J. Cohn for discussing and sharing unpublished results. μ CT scans were performed at the Center for Nanoscale Systems, Harvard University (supported by National Science Foundation award ECS-0335765) and at the Museum of Comparative Zoology. Next-generation sequencing was performed at the HMS Biopolymers Facility and computational analyses were run on the Orchestra Cluster, HMS Research Computing. P.T. was supported by post-doctoral fellowships from the Swiss National Science Foundation, EMBO and the Human Frontiers Science Program. A.C.G. was supported by a post-doctoral fellowship from the Swiss National Science Foundation. This work was supported by National Institutes of Health grant R37-HD032443 to C.J.T.

Author Contributions P.T., J.G. and C.J.T. conceived the project and designed the experiments. P.T. performed most experiments and computational analyses. E.S. prepared CT scans and helped with statistical analyses. T.J.S. helped with CT scans, *Anolis* husbandry and embryo collection. A.C.G. produced lentiviruses and A.C.A. helped with grafting experiments. J.K.H., O.P. and J.G. initiated snake analyses. O.P. contributed snake embryos. J.G. contributed to chick lineage tracing experiments. P.T., J.G. and C.J.T. wrote the paper, with comments from co-authors.

Author Information Sequencing data has been deposited in the Gene Expression Omnibus under accession number GSE60373. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.J.T. (tabin@genetics.med.harvard.edu) or J.G. (jgros@pasteur.fr).

METHODS

Tissue sample collection. All embryos were collected in accordance with the appropriate Institutional Animal Care and Use Committee (IACUC) guidelines. Timed-pregnant CD1 and C57BL/6 females were purchased from Charles River Laboratories. Gravid *Anolis* females were purchased from Candy Quality Reptiles (IACUC #26-11 and #28-14). Anole housing and egg incubation were done as previously described³². To collect early stage pre-oviposition embryos, females were euthanized by intraperitoneal Euthasol injection and eggs were dissected from the oviduct. Snake husbandry and egg collection have been described before³³. Fertilized white leghorn chicken eggs were obtained from Charles River Laboratories and incubated at 38 °C. For staging of mouse embryos, noon on the day of the vaginal plug was considered as E0.5. For chicken and anole embryos, staging was performed according to Hamburger and Hamilton³⁴ or Sanger and colleagues³², respectively. Embryonic tissue was dissected in cold PBS and either fixed in 4% PFA and processed for cryo-embedding or CT scanning, or else directly processed for RNA extraction or stored in RNAlater (Qiagen).

CT scanning and image processing. Embryos were fixed in 4% PFA and stored in 100% ethanol. Staining was done for 48 h in 30% PTA (1% (w/v) phosphotungstic acid in water) and 70% ethanol³⁵. Specimens were rinsed and stored in 70% ethanol until image acquisition. CT scans were carried out on a Bruker Skyscan 1173 or a Nikon (Metris) X-tek HMXST225, at 50–57 kV, 115–145 μ A and 810–1,000 ms exposure time. Voxel sizes ranged from 0.0023 to 0.0056, with 1,500 to 2,400 total projections. Post-processing of scan data was done in VGStudio MAX 2.2 (Volume Graphics). For three-dimensional reconstruction of the cloaca, serial TIFF stacks produced by VGStudio were read into the Imaris software package (Bitplane) and the endodermal epithelium was used as a guide to manually outline the extent of the cloacal volume.

Immunohistochemistry and *in situ* hybridization. Fixed embryos were embedded in 7.5% gelatin/15% sucrose or dehydrated in sucrose gradients and embedded in OCT. Sectioning was performed on a Leica CM3000 cryostat. For immunohistochemistry, sections were incubated with primary antibodies in PBST (PBS/BSA 0.2%, Triton 0.1%, SDS 0.02%) overnight, washed twice for 10 min in PBST and incubated for 1 h with secondary antibodies. To detect genital tubercle and ectopic limb *Isl1* expression, as well as *Lmx1b*, the signal was amplified using the TSA Plus Cy3 kit (Perkin Elmer). Primary antibodies used were anti- β -catenin (BD Biosciences), anti-Shh (Santa Cruz), anti-laminin (Sigma), anti-GFP (Abcam), anti-QCPN, anti-*Lmx1b* and anti-*Isl1* (all Developmental Studies Hybridoma Bank). *In situ* hybridization was performed using standard protocols³⁶. Fluorescent images were acquired on a Zeiss LSM10 inverted confocal microscope. Bright-field images were acquired on a Nikon Eclipse E1000. Whole-mount images were acquired on a Leica MZ FL III. Images were globally processed for colour balance and brightness using Adobe Photoshop.

Lineage tracing analyses. For lentiviral lineage tracing, viral particles harbouring ubiquitously expressing GFP cassettes (UbiC-GFP or hPGK-GFP) were produced by transient transfection in 293T cells as described elsewhere³⁷. Viral particles were then injected into either the coelomic cavity, or the tail bud mesenchyme of mouse E9.5 embryos, chicken HH14 or *Anolis* stage 2–3. For mouse experiments, timed-pregnant CD1 females were anaesthetized using isoflurane and surgery, *in utero* visualization of embryos and virus injection was done as previously outlined³⁸. For chicken embryos, eggs were lowered and windowed and virus injected using a pressure injector. For *Anolis* lineage tracing, we developed a novel whole-embryo *ex ovo* culturing system, using media conditions previously described for squamate organ cultures³⁹. Briefly, we prepared culture dishes with an indentation by pouring 1% Agar Noble (BD Difco) dissolved in culture medium into cell culture dishes containing a modelling clay, egg-shaped casting mould. Once solidified, the mould was removed and stage 2–3 *Anolis* embryos, dissected in 2× PBS, were placed with their yolk intact in the resulting cavity and covered with culture medium. 10% ink in 1× PBS/PenStrep was mouth-pipetted underneath the embryo, for better visualization, and lentiviral particles were injected using a pressure injector. To increase viral infection rate, embryo plates were kept for 12–16 h at 37 °C in a humidified chamber, before switching them to 28 °C. Additional tail bud injections were performed using DiI. Embryos survived for up to 12 days. Only live specimens showing overall normal morphology were considered for further analysis.

To assess GFP⁺ cell contribution, embryos were dissected, fixed in 4% PFA, gelatin-embedded and cryo-sectioned. Sections were stained for GFP and imaged on a Zeiss LSM10 inverted confocal microscope. For quantifications, 4–5 embryos per condition and species were imaged on multiple sections spanning the respective organs. GFP⁺ cell counting was performed in ImageJ, using the 'ITCN' plugin written by T. Kuo (UCSB). A total of 59,331 GFP⁺ cells were counted (mouse: 33,853; chicken: 23,710; anole: 1,768). Counts were averaged over multiple sections and normalized on tissue area measured. The resulting ratios of the two tissues are given as a percentage of total GFP⁺ cells per area.

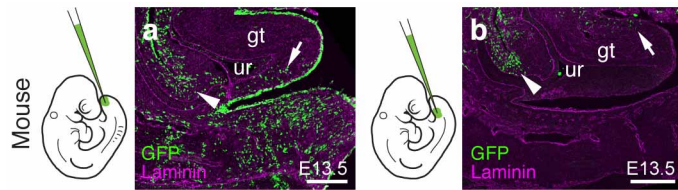
RNA sample preparation and sequencing. Total RNA was extracted from freshly dissected or RNAlater-preserved tissue using the Arcturus PicoPure RNA Isolation Kit (Life Technologies) and enriched for mRNA fraction using MPG mRNA Purification Kit (PureBiotech). Multiplexed RNA-seq libraries were produced using the SPRIA cDNA synthesis kit and the Ovation Ultralow DR Multiplex system (NuGen). Sequencing was done on an Illumina HiSeq 2000, with eight samples multiplexed per lane. Base calling was performed using the Illumina software. Over 561,000,000 50-bp reads were generated, with an average of 23.4 million reads per sample (median = 22,000,000 reads). For all tissues, biological duplicates were sequenced.

Read mapping and transcriptome analyses. Initial read mapping was performed with the RNA-seq unified mapper (RUM)⁴⁰, using mouse NCBI37/mm9 and *Anolis* AnoCar2.0 genome assemblies, with UCSC mm9 refseq and ASU_Acar v2.1 (ref. 41) annotation files, respectively. This resulted in 10,265 orthologous genes between mouse and *Anolis*. An in-house improved annotation for *Anolis*, generated before the publication of Eckalbar *et al.*⁴¹, yielded concordant results in all downstream analyses. To account for species-specific differences in non-uniquely mapping (NU) reads, we redistributed NU reads based on the number of uniquely mapping (U) reads mapping to the respective loci, following a logic outlined before⁴². Multi-dimensional scaling analysis was carried out on normalized read counts using the edgeR bioconductor package⁴³. Genes differentially expressed between early fore- and hindlimb samples were determined in edgeR, and genes showing consistent changes in mouse and *Anolis* were excluded from further analyses, to dampen potential anterior–posterior differences between the organs. For hierarchical clustering, correlation coefficient and principal component analyses, we calculated transcripts per million (TPM)⁴⁴ values, which were then log₂-transformed. Hierarchical clustering was done using the 'pvcust' R package⁴⁵, with 1,000 iterations of multi-scale bootstrap re-sampling, and approximately unbiased (AU) *P* values are provided in the graph. Heat maps of correlation coefficients were plotted with the 'lattice' R package. Principal component analysis was done with the 'prcomp' function in the 'stats' R package and 'GOseq'⁴⁶ was used for GO-term enrichment analysis. Pairwise differential expression analysis between early and late budding stages, in mouse and *Anolis* limbs and genitalia, was done in edgeR, and Venn diagram of genes with an absolute log(fold change) > 1.5 and *P* value < 0.05 was visualized using VennDiagram⁴⁷.

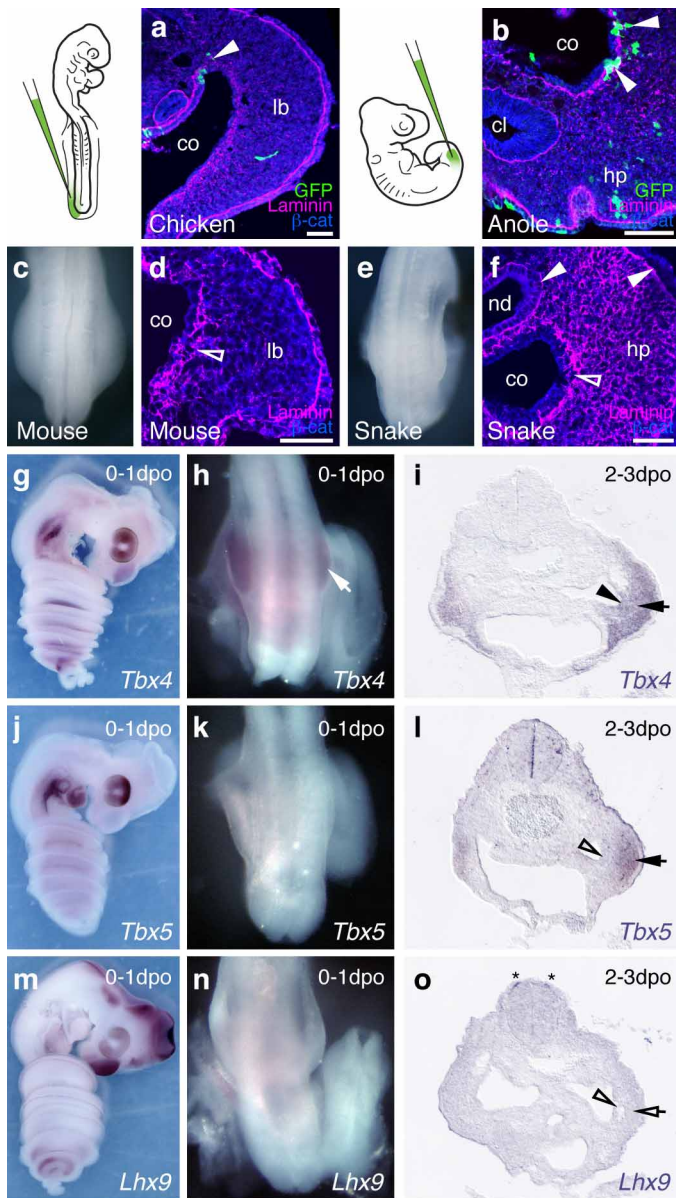
Grafting experiments. For heterotopic, homochronic cloacal grafts, donor and recipient embryos were incubated to reach stage HH17–20. Donor embryos were either GFP-transgenic chicken⁴⁸, purchased from Clemson University, or quail, purchased from Strickland GameBird Farm. Cloacas were dissected in ice-cold PBS and grafted using tungsten needles to a proximal-ventral position, to mimic the squamate configuration, and removed from the apical ectodermal ridge to avoid *SHH*-induced digit duplications¹⁰, or the tail bud in wild-type recipient chicken. Successful grafts were incubated for 1–3 additional days, dissected and screened for the appearance of ectopic outgrowths. Donor versus recipient tissue was discriminated using either GFP or QCPN antibody staining on cryo-sections, or GFP fluorescence for whole-mount embryos. Sham surgery or grafting of GFP-positive limb mesenchyme did not cause any comparable outgrowths. Cloaca-induced outgrowths never stained positive for Alcian blue at later stages, indicating that they were not digit duplications (data not shown). For bead experiments, Affi-Gel Blue Gel beads (150–300 μ m; Bio-Rad) were washed in PBS and incubated for 1–2 h at room temperature, in PBS with recombinant proteins (SHH, FGF2, FGF8; all R&D Systems) at concentrations of 0.1–1 μ g μ l⁻¹. Soaked beads were briefly washed in PBS and grafted to limb and tail buds, as outlined for the cloacal grafts. Control grafts using beads soaked in PBS with bovine serum albumin did not yield any observable outgrowths.

32. Sanger, T. J., Losos, J. B. & Gibson-Brown, J. J. A developmental staging series for the lizard genus *Anolis*: a new system for the integration of evolution, development, and ecology. *J. Morphol.* **269**, 129–137 (2008).
33. Gomez, C. *et al.* Control of segment number in vertebrate embryos. *Nature* **454**, 335–339 (2008).
34. Hamburger, V. & Hamilton, H. L. A series of normal stages in the development of the chick embryo. *J. Morphol.* **88**, 49–92 (1951).
35. Metscher, B. D. MicroCT for developmental biology: a versatile tool for high-contrast 3D imaging at histological resolutions. *Dev. Dyn.* **238**, 632–640 (2009).
36. McGlenn, E. & Mansfield, J. H. Detection of gene expression in mouse embryos and tissue sections. *Methods Mol. Biol.* **770**, 259–292 (2011).
37. Barde, I., Salmon, P. & Trono, D. *Production and Titration of Lentiviral Vectors* (John Wiley & Sons, 2001).
38. Punzo, C. & Cepko, C. L. Ultrasound-guided *in utero* injections allow studies of the development and function of the eye. *Dev. Dyn.* **237**, 1034–1042 (2008).
39. Buchtová, M. *et al.* Initiation and patterning of the snake dentition are dependent on Sonic Hedgehog signaling. *Dev. Biol.* **319**, 132–145 (2008).
40. Grant, G. R. *et al.* Comparative analysis of RNA-Seq alignment algorithms and the RNA-Seq Unified Mapper (RUM). *Bioinformatics* **27**, 2518–2528 (2011).
41. Eckalbar, W. L. *et al.* Genome reannotation of the lizard *Anolis carolinensis* based on 14 adult and embryonic deep transcriptomes. *BMC Genomics* **14**, 49 (2013).

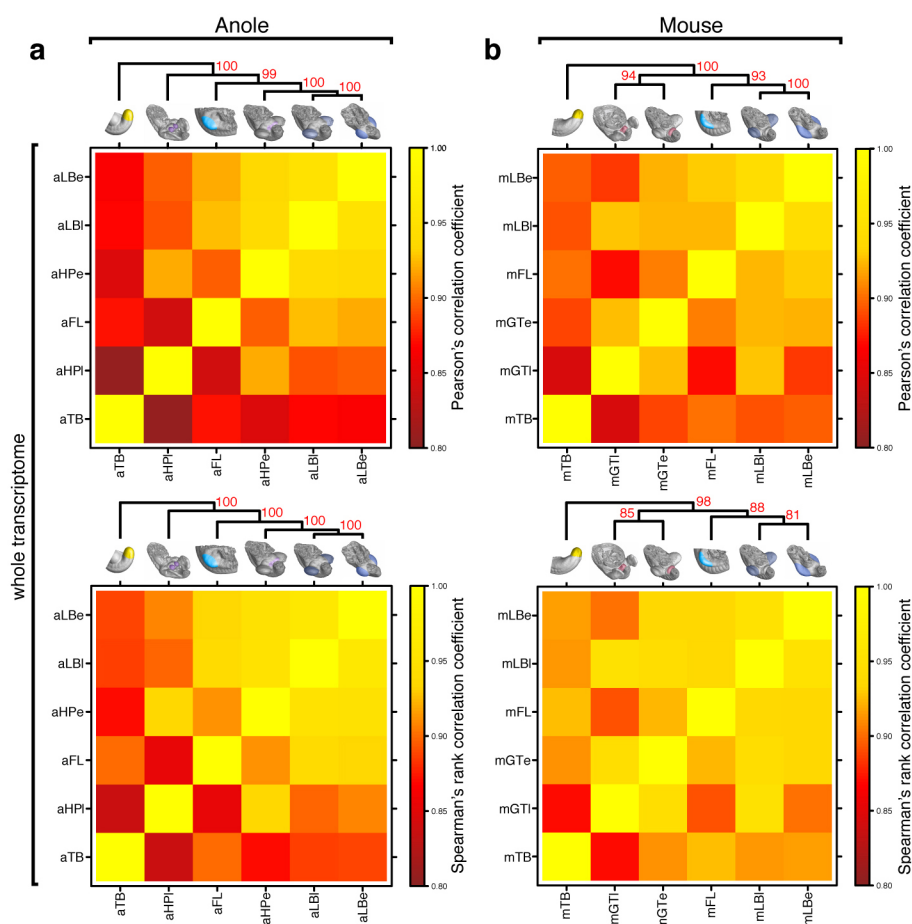
42. Brawand, D. *et al.* The evolution of gene expression levels in mammalian organs. *Nature* **478**, 343–348 (2011).
43. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
44. Wagner, G. P., Koryu, K. & Lynch, V. J. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* **131**, 281–285 (2012).
45. Suzuki, R. & Shimodaira, H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**, 1540–1542 (2006).
46. Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* **11**, R14 (2010).
47. Chen, H. & Boutros, P. C. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12**, 35 (2011).
48. McGrew, M. J. *et al.* Localised axial progenitor cell populations in the avian tail bud are not committed to a posterior Hox identity. *Development* **135**, 2289–2299 (2008).



Extended Data Figure 1 | Two separable ventral cell populations give rise to the murine genital tubercle. **a, b,** Injection into the most distal ventral part of the embryo, the tail bud, marks cells posterior/ventral to the phallic part of the urethra (**a**, arrow; $n = 7$), whereas injection closer to the allantois, into the infra-umbilical mesenchyme, labels cells anterior/dorsal to the phallic part of the urethra (**b**, arrow; $n = 4$). Cells lining the peritoneal cavity are also marked (arrowheads), owing to accidental piercing of the coelom. gt, genital tubercle; ur, urethra. Scale bars, 200 μm .

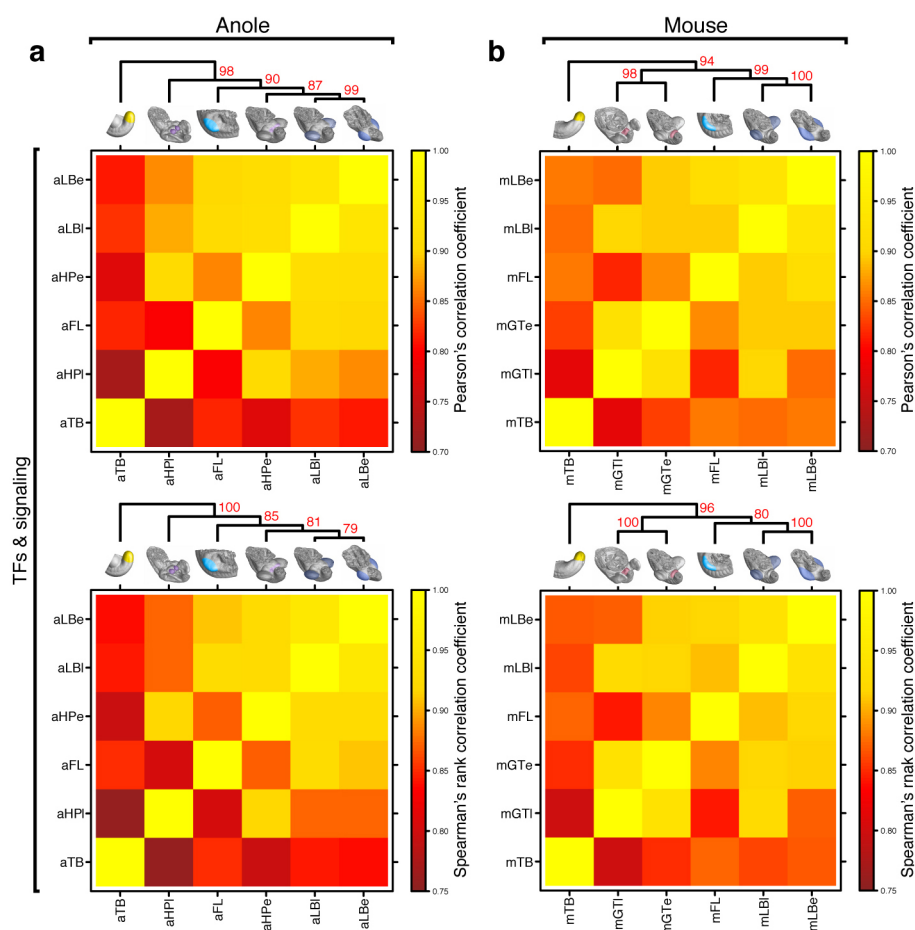


Extended Data Figure 2 | The squamate hemipenis mesenchyme initiates with limb-like cellular dynamics from the coelomic epithelium through an EMT. **a**, Injection of GFP-expressing lentiviruses into the coelom of chicken embryos at HH14 labels cells emerging from the epithelium that contribute to the hindlimb mesenchyme (arrowhead). **b**, In lizards, labelled cells leaving the coelomic epithelium contribute to the hemipenis mesenchyme (arrowheads). **c**, Dorsal view of the hindlimb region of an E10.0 mouse embryo. **d**, Transversal section of a limb bud, showing EMT of the coelomic epithelium (diffuse laminin staining, open arrowhead), as cells contribute to the limb-bud mesenchyme. **e**, Dorsal view of the budding hemipenis of a snake embryo, 1 day after egg deposition. **f**, Transversal section of the hemipenis region. The basement membrane of the coelomic epithelium is breaking down (open arrowhead), while it is intact for both the nephric duct and the surface ectoderm (arrowheads). **g–o**, Expression of genitalia and limb genes during hemipenis initiation. **g–i**, *Tbx4* is expressed early (**h**, arrow) and late during hemipenis initiation, in both the coelomic epithelium (**i**, arrowhead) and the hemipenis mesenchyme (**i**, arrow). **j–l**, *Tbx5* is only expressed later, in the mesenchyme (**l**, arrow), but is absent from the coelomic epithelium (**l**, open arrowhead). **m–o**, Limb marker gene *Lhx9* (see also Fig. 4e) is absent from both epithelium (**o**, open arrowhead) and mesenchyme (**o**, open arrow), but can be detected in dI1 neurons (**o**, asterisk). All gene expression was assessed in at least $n = 3$ samples. cl, cloaca; co, coelom; hp, hemipenis; lb, limb; nd, nephric duct. Scale bars, 50 μ m.



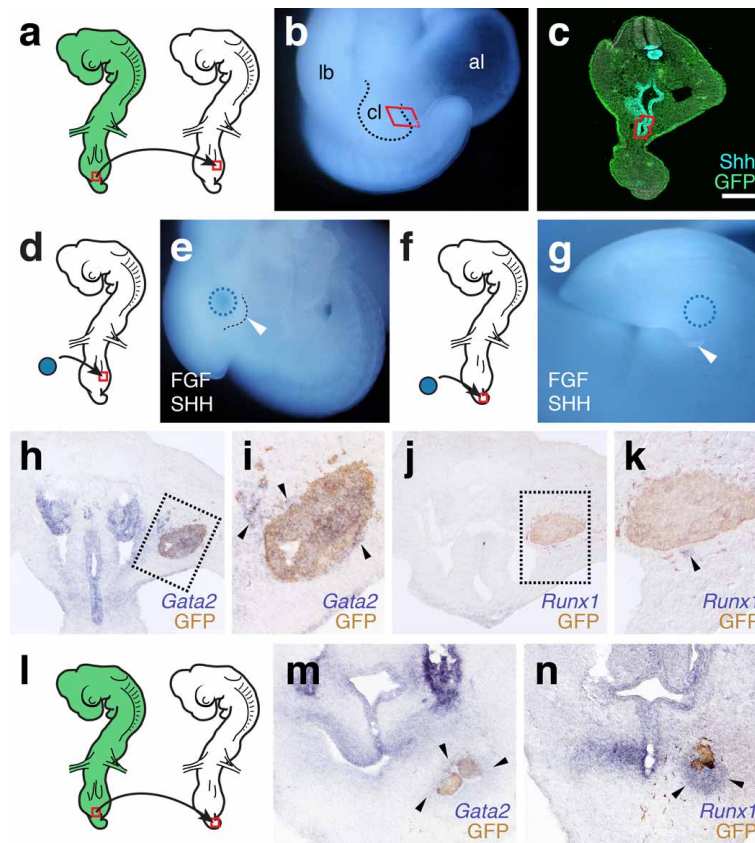
Extended Data Figure 3 | Heat maps of Pearson's and Spearman's rank correlation coefficients and cluster analysis of whole-transcriptome data. **a, b**, Hierarchical clustering on pairwise correlation coefficients for whole-transcriptome data from anole (**a**) and mouse (**b**) samples. Numbers

at nodes represent approximately unbiased *P* values obtained by multiscale bootstrap resampling. Sample identifiers: a, anole; m, mouse; GT, genital tubercle; HP, hemipenis; LB, limb; e, early; l, late.



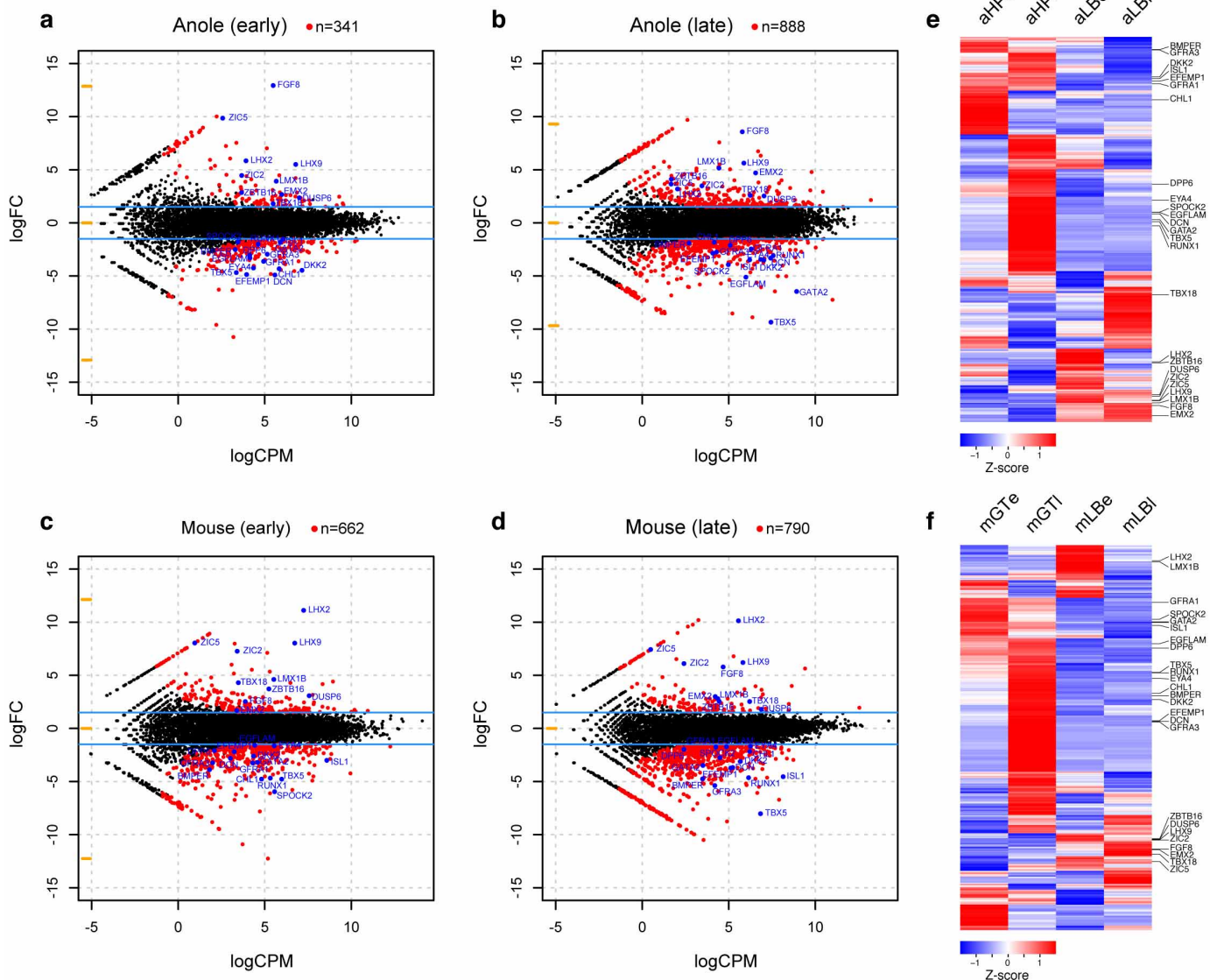
Extended Data Figure 4 | Heat maps of Pearson's and Spearman's rank correlation coefficients and cluster analysis of transcription factor and signalling pathway data. a, b, Hierarchical clustering on pairwise correlation coefficients of transcription factor (TF) and signalling pathway data from

anole (a) and mouse (b) samples. Numbers at nodes represent approximately unbiased *P* values obtained by multiscale bootstrap resampling. Sample identifiers: a, anole; m, mouse; GT, genital tubercle; HP, hemipenis; LB, limb; e, early; l, late.



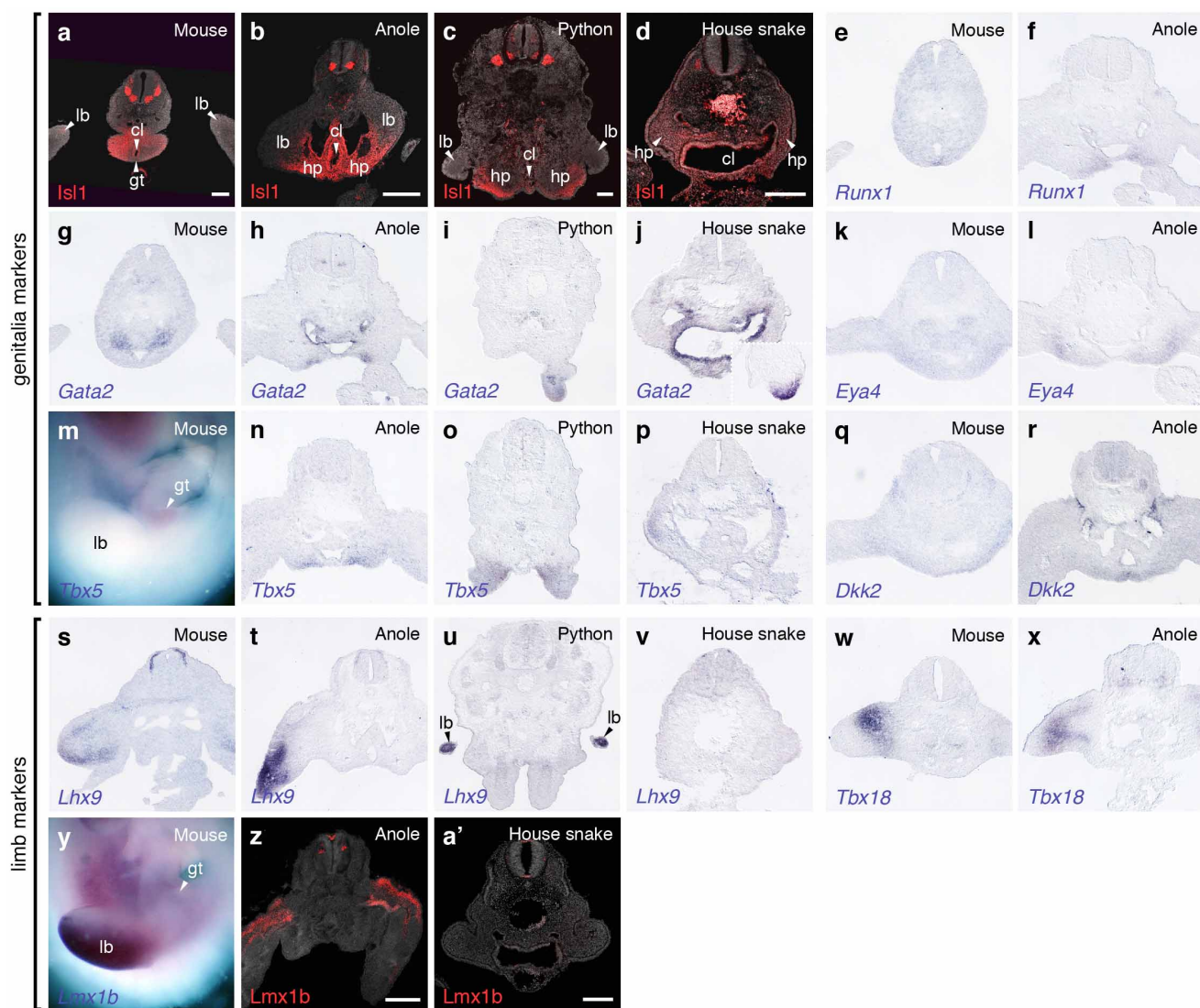
Extended Data Figure 5 | Heterotopic grafting of the cloacal signalling centre leads to ectopic outgrowths and genitalia-like transcriptional changes. **a–c**, Schematics and close-up images of the cloacal grafting procedure. **a**, The cloaca of a stage HH17–19 GFP-transgenic chicken embryo (red rectangle) is transplanted into the proximal-ventral portion of the limb of a wild-type embryo. **b, c**, Only the ventral-most part of the cloaca, including the cloacal membrane, is dissected out (**b**, red box), and subsequently cleared of excess mesenchymal cells attached to the *SHH*-expressing endoderm (**c**, red

outline). **d–g**, Grafting of beads soaked in SHH and FGF can induce ectopic outgrowths on both limbs (**e**; $n = 6/48$) and tail (**g**; $n = 3/31$). **h–k**, Ectopic expression of genital markers *GATA2* (**h, i**, arrowheads) and *RUNX1* (**j, k**, arrowhead) in limb buds, following cloaca-to-limb grafts. **l–n**, Ectopic expression of genital marker *GATA2* (**m**, arrowheads) and *RUNX1* (**n**, arrowheads) in the tail region, following cloaca-to-tail grafts. All gene expression was assessed in at least $n = 3$ samples. al, allantois; cl, cloaca; lb, limb. Scale bar, 200 μ m.



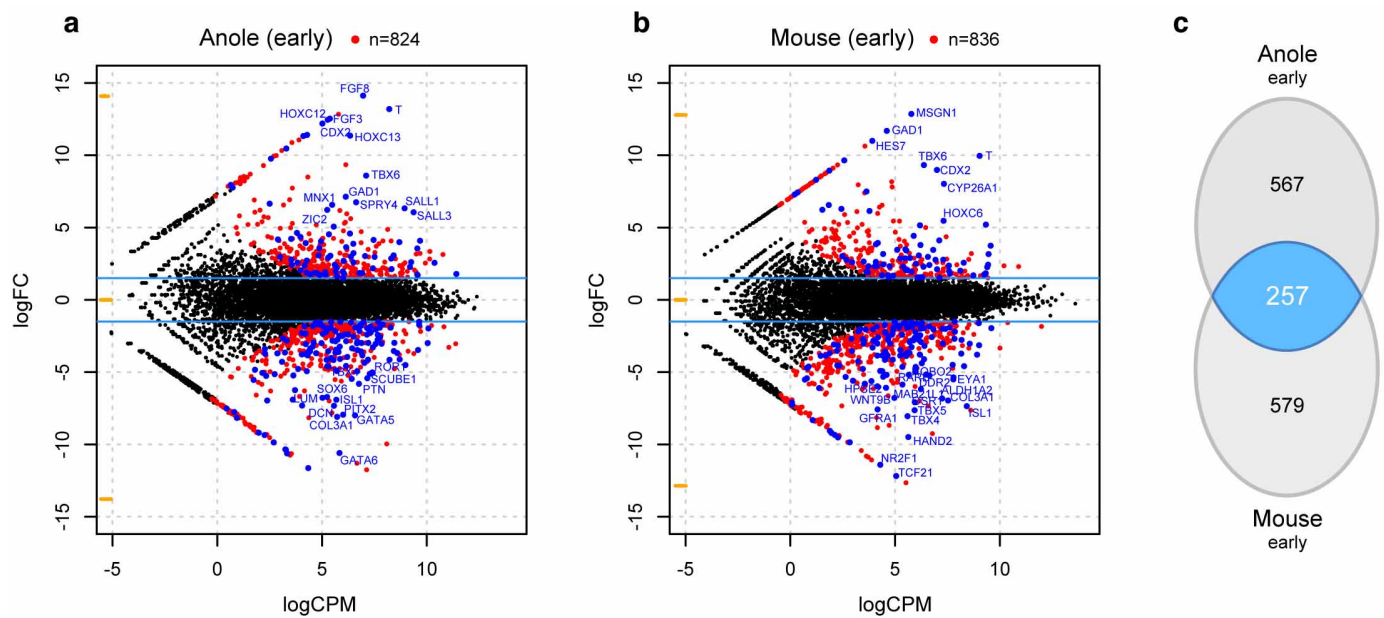
Extended Data Figure 6 | Pairwise differential expression analysis of limb and genitalia transcriptomes. **a–d**, Smear plot visualization of differential expression analyses of early anole (**a**), late anole (**b**), early mouse (**c**) and late mouse (**d**) limb versus genitalia transcriptomes. Genes used for the Venn diagram in Fig. 4d ($|\log_2(\text{fold change})| > 1.5$; P value < 0.05) are highlighted in

red, core 25 marker genes (see Fig. 4e and text) are highlighted and labelled in blue. CPM, counts per million; FC, fold change. **e, f**, Heat map of Z-score-normalized expression values for all genes fulfilling Venn diagram criteria ($n = 2,003$), for anole (**e**) and mouse (**f**) data. Row-based hierarchical clustering was used; core 25 marker genes are indicated on the right.



Extended Data Figure 7 | Comparative marker gene expression analysis in mouse and squamate embryos. a–r, Genitalia markers *Isl1* (a–d), *Runx1* (e, f), *Gata2* (g–j), *Eya4* (k, l), *Tbx5* (m–p) and *Dkk2* (q, r). *Gata2* only becomes visibly expressed at the later stages of house snake hemipenis development

(j, inset). s–a', Limb markers *Lhx9* (s–v), *Tbx18* (w, x) and *Lmx1b* (y–a'). All gene expression was assessed in at least $n = 3$ samples. cl: cloaca; gt: genital tubercle; hp: hemipenis; lb: limb. Scale bar, 200 μ m.



Extended Data Figure 8 | Pairwise differential expression analysis of tail bud and genitalia transcriptomes. **a, b**, Smear plot visualization of differential expression analyses of early anole **(a)** and early mouse **(b)** tail bud versus genitalia transcriptomes. Genes used as input for the Venn diagram in **c** ($|\log_2(\text{fold change})| > 1.5$; $P \text{ value} < 0.05$) are highlighted in red, overlapping

257 marker genes are highlighted in blue. Top 25 genes in the two species, based on logCPM (counts per million) and logFC (fold change), are labelled. **c**, Venn diagram showing overlap of pairwise differential expression analysis results ($\log(\text{fold change}) > 1.5$, $P \text{ value} < 0.05$) of tail bud versus genital tissues for early budding stages in both anole and mouse.

Adenosine activates brown adipose tissue and recruits beige adipocytes via A_{2A} receptors

Thorsten Gnäd¹, Saskia Scheibler^{1,2}, Ivar von Kügelgen¹, Camilla Scheele³, Ana Kilić¹, Anja Glöde¹, Linda S. Hoffmann¹, Laia Reverte-Salisa^{1,2}, Philipp Horn¹, Samet Mutlu¹, Ali El-Tayeb⁴, Mathias Kranz⁵, Winnie Deuther-Conrad⁵, Peter Brust⁵, Martin E. Lidell⁶, Matthias J. Betz⁶, Sven Enerbäck⁶, Jürgen Schrader⁷, Gennady G. Yegutkin⁸, Christa E. Müller^{4,9} & Alexander Pfeifer^{1,9}

Brown adipose tissue (BAT) is specialized in energy expenditure, making it a potential target for anti-obesity therapies^{1–5}. Following exposure to cold, BAT is activated by the sympathetic nervous system with concomitant release of catecholamines and activation of β -adrenergic receptors^{1–5}. Because BAT therapies based on cold exposure or β -adrenergic agonists are clinically not feasible, alternative strategies must be explored. Purinergic co-transmission might be involved in sympathetic control of BAT and previous studies reported inhibitory effects of the purinergic transmitter adenosine in BAT from hamster or rat^{6–8}. However, the role of adenosine in human BAT is unknown. Here we show that adenosine activates human and murine brown adipocytes at low nanomolar concentrations. Adenosine is released in BAT during stimulation of sympathetic nerves as well as from brown adipocytes. The adenosine A_{2A} receptor is the most abundant adenosine receptor

in human and murine BAT. Pharmacological blockade or genetic loss of A_{2A} receptors in mice causes a decrease in BAT-dependent thermogenesis, whereas treatment with A_{2A} agonists significantly increases energy expenditure. Moreover, pharmacological stimulation of A_{2A} receptors or injection of lentiviral vectors expressing the A_{2A} receptor into white fat induces brown-like cells—so-called beige adipocytes. Importantly, mice fed a high-fat diet and treated with an A_{2A} agonist are leaner with improved glucose tolerance. Taken together, our results demonstrate that adenosine–A_{2A} signalling plays an unexpected physiological role in sympathetic BAT activation and protects mice from diet-induced obesity. Those findings reveal new possibilities for developing novel obesity therapies.

The purine nucleoside adenosine is a major precursor and breakdown product of ATP⁹, which acts as a co-transmitter in the sympathetic nerve

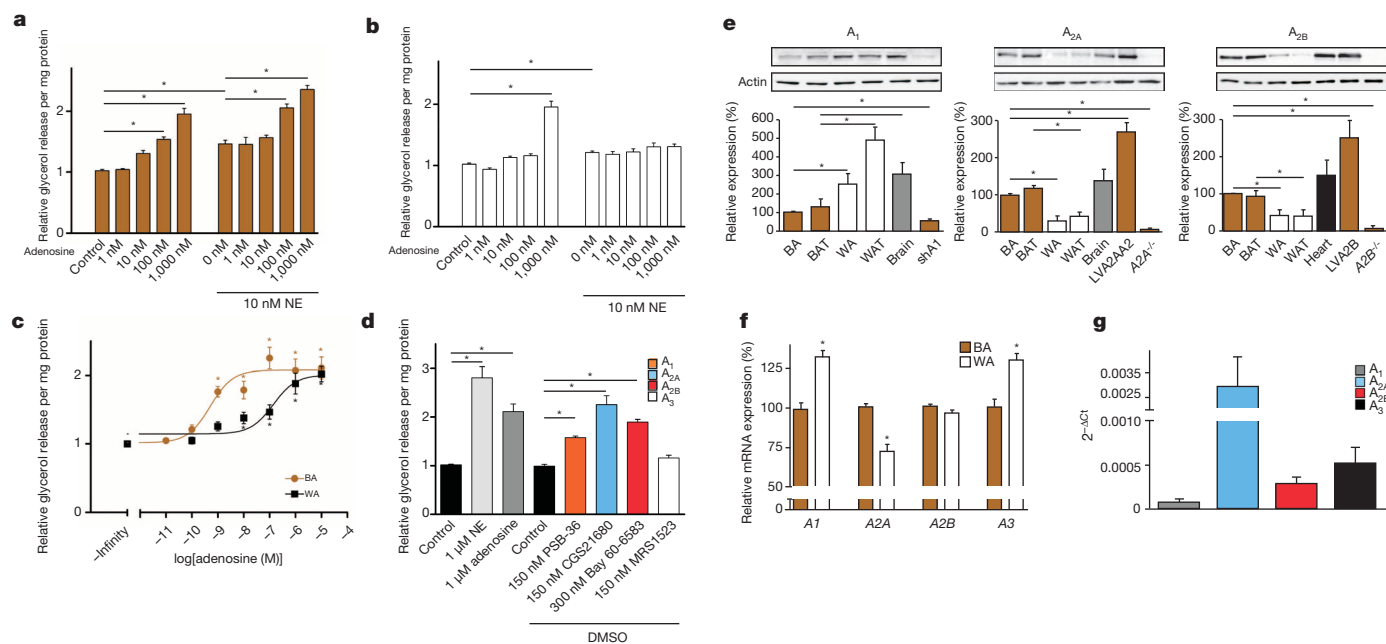


Figure 1 | Activation of human and murine brown adipocytes by adenosine. **a**, **b**, Lipolysis in human brown adipocytes (BA; **a**) and white adipocytes (WA; **b**) treated with adenosine and/or noradrenaline (NE). **c**, Lipolysis in murine brown adipocytes and white adipocytes stimulated with adenosine. **d**, Lipolysis in murine brown adipocytes treated with antagonists for A₁ and A₃ or agonists for A_{2A} and A_{2B}. **e**, Western blotting of adenosine receptors in murine brown adipocytes, white adipocytes, BAT and WAT. LVA2A, LVA2B, A_{2A}^{-/-} or

A_{2B}-overexpressing brown adipocytes by transfection with lentivirus carrying A_{2A} or A_{2B} (also known as *Adora2b*), respectively; shA1, brown adipocytes expressing shRNA-A₁; A_{2A}^{-/-}, A_{2B}^{-/-}, brown adipocytes lacking A_{2A} or A_{2B}, respectively. **f**, Expression of adenosine receptors in human brown adipocytes and white adipocytes. **g**, mRNA expression of adenosine receptor in human BAT normalized to glyceraldehyde 3-phosphate dehydrogenase. *n* = 3; **P* < 0.05. Error bars, s.e.m.

¹Institute of Pharmacology and Toxicology, University Hospital, University of Bonn, 53127 Bonn, Germany. ²Research Training Group 1873, University of Bonn, 53127 Bonn, Germany. ³The Centre of Inflammation and Metabolism and the Centre for Physical Activity Research, Department of Infectious Diseases, Rigshospitalet, 2100 Copenhagen, Denmark. ⁴Pharmaceutical Institute, Pharmaceutical Chemistry I, University of Bonn, 53121 Bonn, Germany. ⁵Helmholtz-Zentrum Dresden - Rossendorf, Institute of Radiopharmaceutical Cancer Research, Research Site Leipzig, 04318 Leipzig, Germany. ⁶Department of Medical and Clinical Genetics, Institute of Biomedicine, The Sahlgrenska Academy, University of Gothenburg, 413 90 Gothenburg, Sweden. ⁷Department for Molecular Cardiology, University of Düsseldorf, 40225 Düsseldorf, Germany. ⁸Medicity Research Laboratory, University of Turku, 20520 Turku, Finland. ⁹Pharma Center, University of Bonn, 53127 Bonn, Germany.

system¹⁰. Adenosine is formed from extracellular ATP by ectonucleotidases¹¹ and alters the function of many cell types¹². The adenosine signal can be transmitted by adenosine A₁ and A₃ receptors through G_i or by A_{2A} and A_{2B} receptors via G_s¹². Adenosine inhibits lipolysis in white adipose tissue (WAT) via the A₁ receptor¹³. In brown adipocytes, lipolysis is essential for activation of thermogenesis: it releases fatty acids and activates the brown fat-specific uncoupling protein 1 (UCP1), thereby converting nutrient energy into heat^{1–5}. Similar to the findings in WAT, adenosine inhibits lipolysis in brown adipocytes from hamsters and rats and reduces the sensitivity to catecholamines^{6–8,14,15}.

In marked contrast to results from these previous studies, adenosine increased lipolysis in a human brown adipocytes cell line (hMADS)^{16,17} (Fig. 1a). Half-maximal activation occurred at 68 nM (Fig. 1a), whereas much higher concentrations of adenosine (1,170 nM) were required to activate primary human white adipocytes (Fig. 1b). Adenosine additively enhanced noradrenaline-induced lipolysis in human brown adipocytes *in vitro*, but not in human white adipocytes (Fig. 1a, b). Moreover, adenosine increased the expression of thermogenic markers in human brown adipocytes and white adipocytes (Extended Data Fig. 1a, b). Importantly, adenosine stimulated an eightfold increase in lipolysis of primary human brown adipocytes derived from supraclavicular BAT¹⁸ with a half-maximal concentration of 3 nM (Extended Data Fig. 1c).

Further detailed analysis required a suitable animal model that—unlike the hamster—mimics the response of human brown adipocytes to adenosine. Treatment of murine brown adipocytes with adenosine increased

lipolysis with a half-maximal effect at 0.7 nM (Fig. 1c). In contrast, 144-fold higher concentrations were required to activate primary murine white adipocytes (Fig. 1c). Genetic knockdown or pharmacological inhibition of A₁ receptors in white adipocytes caused a left-shift of the adenosine concentration-response curve in white adipocytes (Extended Data Fig. 1d, e). Akin to results with human brown adipocytes, adenosine additively enhanced noradrenaline-induced lipolysis in murine brown adipocytes (Extended Data Fig. 1f). Adenosine also induced the expression of thermogenic markers in murine brown adipocytes and white adipocytes (Extended Data Fig. 2a, b). Thus, adenosine activates both murine and human brown adipocytes.

A_{2A} and A_{2B} agonists increased lipolysis by twofold in murine brown adipocytes (Fig. 1d). In contrast, A_{2A} and A_{2B} antagonists reduced adenosine-induced lipolysis and a combination of the two inhibitors completely blocked adenosine effects (Extended Data Fig. 3a). An antagonist for A₃ had no significant effect but blockade of A₁ receptors increased basal lipolysis, suggesting that endogenous adenosine might regulate brown adipocytes (Fig. 1d). Thus, the focus of further analysis was on A₁, A_{2A} and A_{2B} receptors. Adenosine, A_{2A} and A_{2B} agonists, as well as the A₁ antagonist increased cAMP abundance and oxygen consumption (Extended Data Fig. 3b, c). Moreover, adenosine and the A_{2A} agonist increased lipolysis and had an additive effect on noradrenaline effects in BAT explants (Extended Data Fig. 3d).

Western blotting revealed several-fold higher levels of A_{2A} and A_{2B} protein in mature brown adipocytes than in white adipocytes, whereas

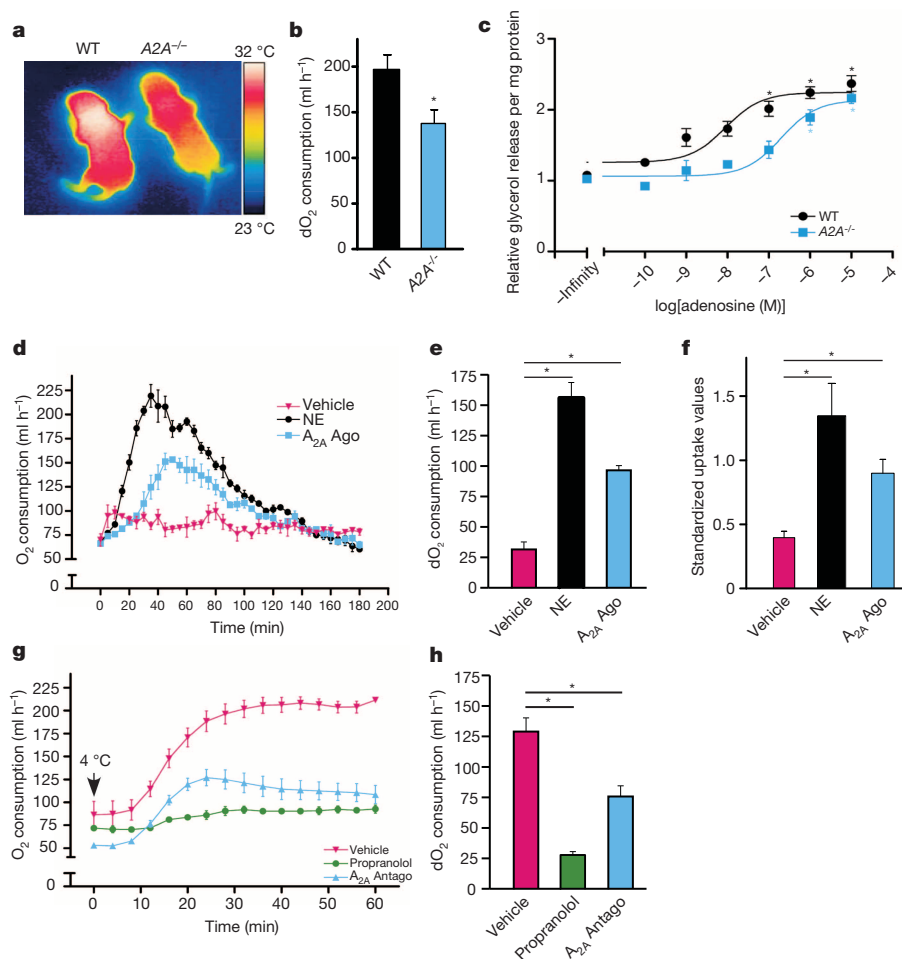


Figure 2 | BAT activation by A_{2A} receptors. **a**, Representative thermographic image of wild-type (WT) and A_{2A}^{-/-} pups. **b**, Relative increase of oxygen consumption of adult A_{2A}^{-/-} and WT mice exposed to 4 °C. **c**, Adenosine-induced lipolysis in WT or A_{2A}^{-/-} brown adipocytes. **d**, **e**, Oxygen consumption in mice injected with A_{2A} agonist (CGS21680), noradrenaline or

vehicle. **f**, PET/MRI analysis of FDG uptake by interscapular BAT in mice treated with noradrenaline or A_{2A} agonist (PSB-0777). **g**, **h**, Oxygen consumption of mice exposed to 4 °C and injected with A_{2A} antagonist or propranolol. *n* = 3 animals and independent cell cultures were analysed; **P* < 0.05. Error bars, s.e.m.

A₁ was the predominant adenosine receptor in murine white adipocytes (Fig. 1e). Differentiation of murine brown adipocytes was associated with an approximately 60-fold increase in expression of A_{2A} receptors (Extended Data Fig. 3e). By contrast, A₁ was more highly expressed during differentiation of white adipocytes as compared to brown adipocytes (Extended Data Fig. 3e), while hamster BAT has similar levels of A₁ and A_{2A} receptors, and undetectable expression of A_{2B} receptors (Extended Data Fig. 3f). Moreover, expression of A_{2A} receptors was significantly higher in human brown adipocytes than in human white adipocytes, whereas the A₁ receptor expression was higher in white adipocytes (Fig. 1f). In human BAT¹⁹, A_{2A} receptors are the most highly expressed adenosine receptors (Fig. 1g). Together, these data show that adenosine receptors are differentially expressed in WAT and BAT as well as in different species, which probably explains why adenosine can have opposing effects in distinct species and in the two types of adipose tissue.

Next, genetic and pharmacological tools were used to define the *in vivo* role of adenosine/A_{2A} signalling. BAT-derived thermogenesis of A_{2A}-deficient (A_{2A}^{-/-}; also known as *Adora2a*^{-/-}) new-born mice²⁰ was significantly reduced compared to wild-type littermates (Fig. 2a and Extended Data Fig. 4a). Indirect calorimetry revealed a ~30% reduction in oxygen consumption of cold-exposed adult A_{2A}^{-/-} mice (Fig. 2b). Oxygen consumption at thermoneutrality, body weight and locomotor activity of A_{2A}^{-/-} mice were not different (Extended Data Fig. 4b–e). Moreover, the histological appearance of A_{2A}^{-/-} BAT was not different from wild type (Extended Data Fig. 4f). Although differentiation of brown pre-adipocytes isolated from A_{2A}^{-/-} mice was not altered (Extended Data Fig. 4g–i), A_{2A}^{-/-} brown adipocytes were more than 20-fold less sensitive to adenosine than wild-type cells (Fig. 2c) and had a similar sensitivity to adenosine as wild-type white adipocytes (see Fig. 1c). The adenosine- and A_{2A}-agonist-induced increase in thermogenic markers was blunted in brown adipocytes and white adipocytes of A_{2A}^{-/-} mice (Extended Data Fig. 2a, b). Furthermore, adenosine and A_{2A} agonist-induced respiration and lipolysis were blunted in the absence of A_{2A} receptors (Extended Data Fig. 4j and Extended Data Fig. 5a–d). Together, these data indicate that adenosine signalling via A_{2A} is required for full physiological activation of BAT.

Injection of A_{2A} agonists strongly increased whole-body oxygen consumption, reaching 70% of the maximal effect provoked by noradrenaline without altering locomotor activity (Fig. 2d, e and Extended Data Fig. 6a–c). Blocking β -adrenergic receptors with propranolol before A_{2A} activation had no significant effect, indicating that the A_{2A} agonist effects are not mediated by β -adrenergic signalling (Extended Data Fig. 6d, e). Furthermore, stimulation with noradrenaline or A_{2A} agonist caused a significantly higher uptake of [¹⁸F]fluorodeoxyglucose (FDG) compared to vehicle treatment into murine BAT as measured with positron emission tomography/magnetic resonance imaging (PET/MRI; Fig. 2f and Extended Data Fig. 6f, g). To study the role of A_{2A} in BAT activation by a physiological stimulus, mice were exposed to cold. A_{2A} receptor expression was increased in cold-exposed mice as well as in brown adipocytes in response to noradrenaline or cAMP (Extended Data Fig. 6h, i). Blockade of A_{2A} receptors significantly diminished cold-induced oxygen consumption (Fig. 2g, h) without affecting locomotion nor the abundance of noradrenaline in BAT (Extended Data Fig. 6j, k); the latter indicates that the effects of the A_{2A} antagonist are not caused by changes in the sympathetic tone to BAT.

Adenosine might be released by two major mechanisms in BAT: breakdown of ATP released from sympathetic nerves¹¹ and/or autocrine/paracrine efflux from brown adipocytes⁶. To induce neurotransmitter release from sympathetic nerve terminals within BAT, electrical field stimulation (EFS) was used. EFS evoked the release of endogenous noradrenaline and ATP (Fig. 3a, b). In parallel, EFS induced a more than sevenfold increase in adenosine concentration compared to unstimulated BAT (Fig. 3c). All of these EFS effects were abolished after blocking action potentials by tetrodotoxin (Fig. 3a–c).

Although BAT from mice deficient in CD73, the ecto-5'-nucleotidase that produces adenosine from extracellular nucleotides¹¹, exhibited lower

basal adenosine levels, the EFS-induced increase in adenosine occurred also in the absence of CD73 (Extended Data Fig. 7a, b). Moreover, the α -adrenergic blocker phenoxybenzamine failed to alter the stimulation-evoked outflow of ATP and adenosine from BAT (Extended Data Fig. 7c, d).

In addition to adenosine release by EFS, treatment of BAT and brown adipocytes with noradrenaline significantly increased adenosine concentrations, which was abolished after treatment with propranolol (Fig. 3d, e). ATP concentrations were not affected by noradrenaline (Fig. 3f and Extended Data Fig. 7e). Thus, there appears to be a cross-talk between adrenergic and purinergic signalling that enhances BAT activation.

Next, the potential of adenosine/A_{2A} signalling to counteract diet-induced obesity was assessed. Mice on high-fat diet (HFD) treated with an A_{2A} agonist exhibited a significant reduction in body weight as well as a 26% reduction in relative fat mass and a 13% increase in lean mass (Fig. 4a, b). The weights of inguinal WAT (iWAT) and gonadal fat (gWAT) depots were reduced by 48% and 71%, respectively, in A_{2A} agonist-treated mice (Extended Data Fig. 8a). Importantly, mice treated with the A_{2A} agonist exhibited increased basal as well as noradrenaline-stimulated, maximal oxygen consumption (Fig. 4c, d) in the absence of altered food intake or locomotor activity (Extended Data Fig. 8b–d). A_{2A} agonist treatment also improved glucose tolerance (Fig. 4e and Extended Data Fig. 8e). Moreover, mice treated with the A_{2A} agonist exhibited increased expression of thermogenic markers in BAT and WAT with a more than sevenfold increase of UCP1 in WAT (Extended Data Fig. 8f, g). The abundance of noradrenaline and of alternatively activated macrophages were not increased in the adipose tissues of animals treated with the A_{2A} agonist, thus ruling out increased sympathetic input or accumulation of macrophages producing catecholamines (Extended Data Fig. 8h, i).

Browning of WAT—the appearance of beige cells—protects mice from diet-induced obesity and can be induced by β -adrenergic agonists and several other stimuli^{3,5,21}. The two brown adipocyte markers PPAR γ coactivator 1 α (PGC1 α) and UCP1 were increased in WAT of mice

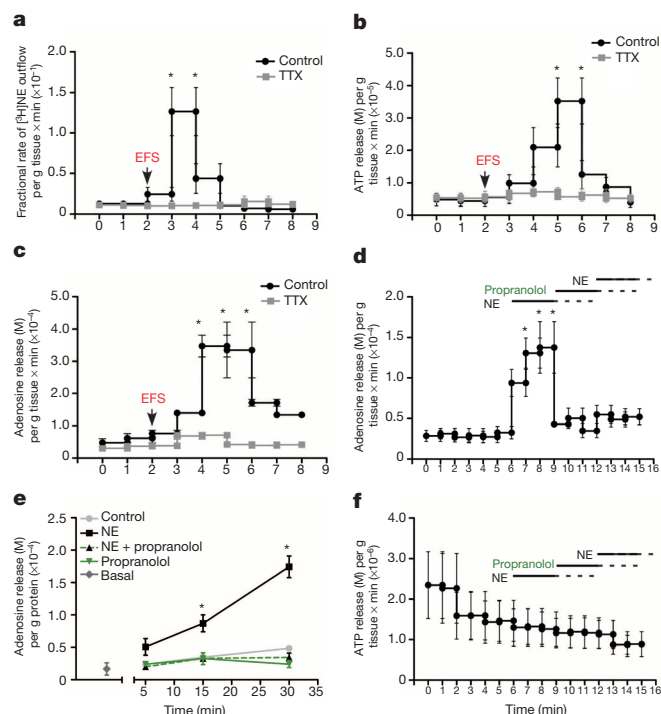


Figure 3 | Adenosine release in BAT. a–c, Release of noradrenaline (a), ATP (b) and adenosine (c) in murine BAT subjected to EFS with and without tetrodotoxin (TTX). d, e, Release of adenosine from BAT (d) or brown adipocytes (e) after stimulation with noradrenaline in presence or absence of propranolol. f, ATP release from BAT treated with noradrenaline in presence or absence of propranolol. $n = 3$ for all experiments; $*P < 0.05$. Error bars, s.e.m.

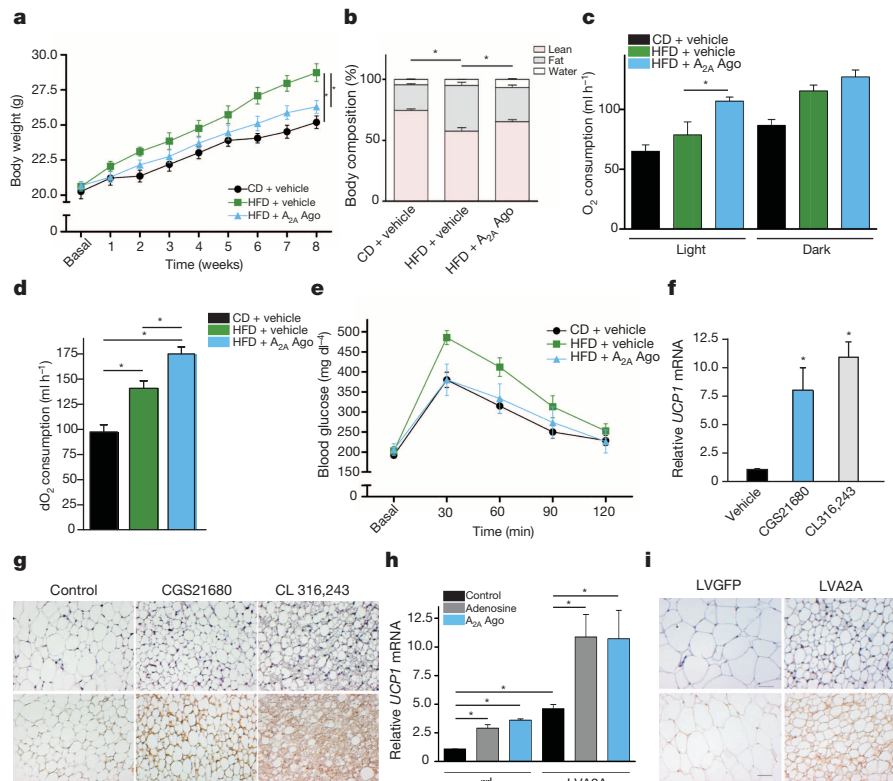


Figure 4 | A_{2A} stimulation prevents diet-induced obesity and induces browning. **a, b**, Body weight (**a**) and body composition (**b**) of mice fed either control (CD) or HFD and treated with an A_{2A} agonist (PSB-0777). **c, d**, Basal (**c**) and noradrenaline-induced (**d**) relative oxygen consumption. **e**, Glucose tolerance test. **f**, UCP1 expression in iWAT of mice treated with CGS21680 or CL316,243 (β_3 agonist) for 10 days. **g**, Representative haematoxylin and eosin

treated for 10 days with an A_{2A} agonist concomitantly with a reduction of adipocyte size and the presence of multilocular, UCP1-positive beige cells in the A_{2A} agonist-treated animals (Fig. 4f, g and Extended Data Fig. 9a), showing that selective activation of A_{2A} receptors in WAT can induce browning.

Because A_{2A} levels are significantly lower in white adipocytes than in brown adipocytes, lentiviral vectors carrying A_{2A} (LVA2A) were used to increase A_{2A} expression in white adipocytes (Extended Data Fig. 9b). LVA2A significantly increased expression of thermogenic markers and lipolysis, which was abolished by adenosine deaminase or an A_{2A} antagonist, thus showing that endogenous adenosine can activate A_{2A} receptors in white adipocytes (Fig. 4h and Extended Data Fig. 9c, d).

To determine whether endogenous adenosine and A_{2A} overexpression can be used to induce browning *in vivo*, LVA2A were injected into iWAT, which possesses a high capacity for browning³ of mice on HFD. Gene transfer of A_{2A} resulted in decreased adipocyte hypertrophy and reduced expression of inflammatory cytokines (Fig. 4i and Extended Data Fig. 9e–g). Moreover, LVA2A-injected WAT exhibited browning with multilocular adipocytes and increased BAT marker expression (Fig. 4i and Extended Data Fig. 9h). Thus, browning of WAT can be induced by adenosine signalling, either by selective stimulation of A_{2A} receptors or by endogenous adenosine after augmenting the abundance of A_{2A}.

In conclusion, we identify signalling by adenosine and in particular by A_{2A} receptors as physiological mechanisms for activation of human and murine BAT. Adenosine acts as an additive or synergistic co-transmitter together with noradrenaline in BAT. Pharmacological A_{2A} receptor stimulation or lentiviral A_{2A} overexpression improves obesity-caused changes and induces browning of WAT. In the light of the world-wide obesity pandemic²², activators of BAT may be potential drug targets for anti-obesity therapies and as shown here, adenosine is a previously unappreciated activator of BAT.

(top) or UCP1 staining (bottom) of iWAT. **h**, UCP1 expression in murine white adipocytes infected with control virus (rrl) or lentivirus carrying A_{2A} (LVA2A) and treated with adenosine (100 nM) or CGS21680 (150 nM). **i**, Representative haematoxylin and eosin (top) or UCP1 staining (bottom) of iWAT injected with lentivirus carrying the green fluorescent protein gene (LVGFP) or LVA2A. $n = 4$ (**a–f**); $n = 3$ (**h**); * $P < 0.05$. Error bars, s.e.m.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 May; accepted 1 September 2014.

Published online 15 October 2014.

- Cannon, B. & Nedergaard, J. Brown adipose tissue: function and physiological significance. *Physiol. Rev.* **84**, 277–359 (2004).
- Kajimura, S. & Saito, M. A new era in brown adipose tissue biology: molecular control of brown fat development and energy homeostasis. *Annu. Rev. Physiol.* **76**, 225–249 (2013).
- Harms, M. & Seale, P. Brown and beige fat: development, function and therapeutic potential. *Nature Med.* **19**, 1252–1263 (2013).
- Bartelt, A. & Heeren, J. Adipose tissue browning and metabolic health. *Nature Rev. Endocrinol.* **10**, 24–36 (2014).
- Rosen, E. D. & Spiegelman, B. M. What we talk about when we talk about fat. *Cell* **156**, 20–44 (2014).
- Schimmel, R. J. & McCarthy, L. Role of adenosine as an endogenous regulator of respiration in hamster brown adipocytes. *Am. J. Physiol.* **246**, C301–C307 (1984).
- Szillat, D. & Bukowiecki, L. J. Control of brown adipose tissue lipolysis and respiration by adenosine. *Am. J. Physiol.* **245**, E555–E559 (1983).
- Woodward, J. A. & Saggerson, E. D. Effect of adenosine deaminase, N⁶-phenylisopropyladenosine and hypothyroidism on the responsiveness of rat brown adipocytes to noradrenaline. *Biochem. J.* **238**, 395–403 (1986).
- Abbracchio, M. P., Burnstock, G., Verkhratsky, A. & Zimmermann, H. Purinergic signalling in the nervous system: an overview. *Trends Neurosci.* **32**, 19–29 (2009).
- Gourine, A. V., Wood, J. D. & Burnstock, G. Purinergic signalling in autonomic control. *Trends Neurosci.* **32**, 241–248 (2009).
- Zimmermann, H., Zebisch, M. & Sträter, N. Cellular function and molecular structure of ecto-nucleotidases. *Purinergic Signal.* **8**, 437–502 (2012).
- Fredholm, B. B., Ijzerman, A. P., Jacobson, K. A., Linden, J. & Müller, C. E. International Union of Basic and Clinical Pharmacology. LXXXI. Nomenclature and classification of adenosine receptors—an update. *Pharmacol. Rev.* **63**, 1–34 (2011).
- Johansson, S. M., Lindgren, E., Yang, J. N., Herling, A. W. & Fredholm, B. B. Adenosine A₁ receptors regulate lipolysis and lipogenesis in mouse adipose tissue—interactions with insulin. *Eur. J. Pharmacol.* **597**, 92–101 (2008).

14. McMahon, K. K. & Schimmel, R. J. Apparent absence of alpha-2 adrenergic receptors from hamster brown adipocytes. *Life Sci.* **30**, 1185–1192 (1982).
15. Unelius, L., Mohell, N. & Nedergaard, J. Cold acclimation induces desensitization to adenosine in brown fat cells without changing receptor binding. *Am. J. Physiol.* **258**, C818–C826 (1990).
16. Rodriguez, A. M. *et al.* Transplantation of a multipotent cell population from human adipose tissue induces dystrophin expression in the immunocompetent mdx mouse. *J. Exp. Med.* **201**, 1397–1405 (2005).
17. Bordinchia, M. *et al.* Cardiac natriuretic peptides act via p38 MAPK to induce the brown fat thermogenic program in mouse and human adipocytes. *J. Clin. Invest.* **122**, 1022–1036 (2012).
18. Jespersen, N. Z. *et al.* A classical brown adipose tissue mRNA signature partly overlaps with brite in the supraclavicular region of adult humans. *Cell Metab.* **17**, 798–805 (2013).
19. Lidell, M. E. *et al.* Evidence for two types of brown adipose tissue in humans. *Nature Med.* **19**, 631–634 (2013).
20. Chen, J. F. *et al.* A_{2A} adenosine receptor deficiency attenuates brain injury induced by transient focal ischemia in mice. *J. Neurosci.* **19**, 9192–9200 (1999).
21. Frontini, A. & Cinti, S. Distribution and development of brown adipocytes in the murine and human adipose organ. *Cell Metab.* **11**, 253–256 (2010).
22. Speakman, J. R. & O'Rahilly, S. Fat: an evolving issue. *Dis. Model. Mech.* **5**, 569–573 (2012).

Acknowledgements We thank D. Hass, S. Kipschull, N. Galicki James, B. Steckel, J. Müllich, M. Schneider and G. Petersson for technical assistance. We thank M. Idzko for A_{2A}^{−/−} mice. A.P. was supported by the Deutsche Forschungsgemeinschaft (DFG); S.S. and L.R.-S. were supported by the DFG-funded Research Training Group 1873 “Pharmacology of 7TM-receptors and downstream signalling pathways”; C. S. was supported by the Lundbeck Foundation and Novo Nordisk A/S. CIM is supported by the Danish National Research Foundation (DNRF55) and CFAS is supported by Trygfondene. A.K. and L.S.H. were supported by BONFOR. A.G. was supported by the Bonn International Graduate School DrugS and the Federal Ministry of Education and

research. M.E.L. was supported by the Swedish Research Council (2013-4466). S.E. was supported by the Swedish Research Council (2010-3281, 2012-1652), The Knut and Alice Wallenberg Foundation, Sahlgrenska's University Hospital (LUA-ALF), European Union grants (HEALTH-F2-2011-278373; DIABAT), the Inga Britt and Arne Lundgren Foundation, the Söderberg Foundation, and the King Gustaf V and Queen Victoria Freemason Foundation.

Author Contributions T.G. designed and performed most experiments, analysed the data and wrote the manuscript. S.S. performed EFS experiments. I.v.K. set up and designed EFS experiments and analysed the data. C.S. differentiated primary brown adipocytes and helped with lipolysis experiments. A.K. performed short term browning experiments and participated in HFD studies. A.G. helped with hMADS and hWA experiments. L.S.H. helped with lentiviral injection experiments and iGTT. L.R.-S. performed *ex vivo* lipolysis experiments and analysed the data. P.H. helped with lipolysis and cAMP experiments. S.M. helped with murine brown adipocytes lipolysis experiments. A.E.-T. synthesized PSB-0777 and determined PSB-0777 levels in blood and brain samples. M.E.L. performed qPCR and analysed human PCR data. M.J.B. collected human BAT samples and performed qPCR. S.E. analysed human PCR data and revised the manuscript. J.S. provided HPLC data and revised the manuscript. G.G.Y. set up extracellular purine measurements and revised the manuscript. C.E.M. suggested and provided compounds, discussed data, supervised determination of PSB-0777 levels and revised the manuscript. A.P. supervised all experiments and wrote the manuscript. M.K. participated in the design and carried out the PET/MRI studies, was responsible for the PET data analysis. W.D.-C. participated in the design and performance of the PET/MRI studies, supported the PET data analysis. P.B. participated in the design and data analysis of the PET/MRI studies.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.P. (alexander.pfeifer@uni-bonn.de).

METHODS

Materials. Antibodies against A₁ (H-40), A_{2A} (7F6-G5-A2), A_{2B} (H-40), αP2 and PPARγ were purchased from Santa Cruz. Antibodies against UCP1 and actin were from Sigma and against tubulin from Dianova. TTX, adenosine, propranolol and noradrenaline were bought from Sigma-Aldrich. PSB-36, PSB-603, PSB-0777, MSX-2 and MSX-3 were provided by C.E. Müller. CGS21680, Bay 60-6583, MRS1523 and CL 316,243 were obtained from Tocris. Noradrenaline/Arterenol was purchased from Sanofi Aventis.

Primary human and murine adipocyte culture. Stromal vascular cells from human supraclavicular adipose tissue biopsies¹⁸ and mouse intrascapular BAT²³ were isolated and differentiated as described previously. hMADS¹⁶ were provided by the laboratory of C. Dani (University of Nice Sophia Antipolis). Human white adipocytes were obtained from Promo Cell and differentiated according to manufacturer's instructions. Murine pre-adipocytes were isolated from iWAT and differentiated to mature white adipocytes as described²³ and UCP1 and PGC1α levels were analysed after 8 h stimulation of mature cells.

Measurement of endogenous respiration. Mature brown adipocytes were treated as indicated 30 min before oxygraphic measurements (Oxygraph 2K; Oroboros Instruments). Cell layer was transferred to the oxygraph chamber containing 2 ml incubation medium (0.5 mM EGTA, 3 mM MgCl₂ 6H₂O, 60 mM K-lactobionate, 20 mM taurine, 10 mM KH₂PO₄, 20 mM HEPES, 110 mM sucrose and 1 g l⁻¹ BSA, pH 7.1) with 25 mg ml⁻¹ digitonin. *In vitro* respiration levels of the adipocytes were recorded when reaching a steady state. Respiration rates were normalized to total protein content.

Lipolysis assay. Differentiated adipocytes were washed twice with lipolysis medium (Life Technologies) supplemented with 2% w/v fatty acid-free BSA (Sigma-Aldrich) followed by incubation with lipolysis medium containing indicated substances at 37 °C and 5% CO₂ for 2 h. For the *ex vivo* lipolysis, BAT from newborn mice was isolated and tissue explants were minced and incubated with lipolysis medium containing the indicated substances at 37 °C and 5% CO₂ for 2 h.

The media were collected and after an incubation of 5 min at 37 °C with free glycerol reagent (Sigma-Aldrich) absorption was measured at 540 nm. Glycerol release was calculated with glycerol standard (Sigma-Aldrich) and normalized to protein content.

Western blot and quantification. Protein amount from all samples was quantified using Bradford assay followed by concentration normalization before western blot experiments. Western blot was carried out following standard procedures and band intensity was quantified using ImageJ. All data were normalized to background and loading controls.

Electrical field stimulation. EFS (1,000 pulses at 10 Hz) was performed as described previously²⁴. In brief, noradrenaline release was measured by tritium outflow from BAT after pre-incubation with ³H-labelled noradrenaline (Perkin Elmer) for 30 min. ATP was measured using luciferase bioluminescence assay (ATP Bioluminescence assay Kit HS II, Roche). Results were normalized to wet tissue weight. Action potentials were blocked with 1 μM TTX.

RNA isolation and qPCR. Total RNA was isolated from cells or tissues using TRIzol (Invitrogen). Reverse transcription was performed using Transcriptor First Strand Synthesis Kit (Roche). qPCR was performed with SYBR-Green (Roche) using a HT7900 instrument (Applied Biosystems). Fold changes were calculated using relative quantification methods with hGAPDH (human glyceraldehyde 3-phosphate dehydrogenase) or mHPRT (murine hypoxanthine guanine phosphoribosyl transferase) serving as internal control. For expression analysis of human retroperitoneal BAT, RNA was isolated and cDNA synthesized as described before²⁵; qPCR was performed on a ViiA7 instrument (Applied Biosystems) using Power SYBR Green PCR master mix from Life Technologies.

Measurement of extracellular purines. Adenosine was assayed using specific mixtures of enzymes converting purines into uric acid and H₂O₂, followed by fluorometric detection of the generated H₂O₂ by Amplex Red Reagent (emission/excitation wavelengths 545/590 nm)²⁶. All measurements were performed on Perkin Elmer Enspire microplate reader, and calibration curves were generated for each experiment using identical coupled reactions with serial dilutions of exogenous purine standards. Results were normalized to wet tissue weight or protein content. BAT or brown adipocytes were stimulated with noradrenaline (1 μM) in presence or absence of propranolol (2 μM).

Adenosine concentrations were additionally measured using a WATERS high-performance liquid chromatography (HPLC) system equipped with an analytical Hypersil BDS column (4.6 × 150 cm, Thermo Scientific, Waltham, USA), as described previously²⁷. Identification and quantification of adenosine peaks was done by comparison to retention times of known standards and peak integration and normalization.

cAMP determination. cAMP levels were measured by EIA (Cayman Chemical) following the manufacturer's instructions. Results were normalized to protein content.

Noradrenaline measurement. Noradrenaline levels were determined by ELISA (CatCombi) with freshly isolated adipose tissue following manufacturer's instructions. Results were normalized to wet tissue weight.

Plasmids and viral infection. Lentiviral vectors were constructed by cloning murine A2A or eGFP complementary DNA into the p156rrlCMV vector (LVA2A; LVGFP). The control vector contained no transgene (rrl). Lentivirus production was performed as previously described²³. For *in vitro* A2A overexpression 50 ng of virus particles were used.

In vivo analysis. Animals. WT male C57Bl/6 mice at indicated ages and Golden Syrian hamsters were purchased from Charles River. A_{2A} knockout animals²⁰ were purchased from The Jackson Laboratory (Strain C; 129-Adora2atm1fc/J).

HFD (60% energy from fat) and control diet was purchased from Ssniff. Mice were maintained on a daily cycle of 12 h light (06:00 to 18:00) and 12 h darkness (18:00 to 06:00), at 24 ± 1 °C, and were allowed free access to chow and water.

For the diet-induced obesity study, mice were randomized for weight and were injected intraperitoneally once daily with A_{2A} agonist PSB-0777 (1 mg per kg body weight), a polar substance that does not cross the blood-brain barrier (as determined by liquid chromatography/mass spectrometry of brain samples; data not shown), for eight weeks.

Lentiviral vectors coding for A_{2A} (LVA2A) were used to overexpress A_{2A} in WAT; lentiviral vectors coding for GFP were used as control. 10-week old male C57Bl/6 mice randomized for weight were anaesthetized with isoflurane, a short incision was made in the flank and viral particles (350 ng lentiviral particles/fat pad (p24 antigen Elisa, ZeptoMetrix)) were directly injected into the iWAT depots and animals were put on HFD for 6 weeks.

For short-term browning of WAT, 10-week-old male C57Bl/6 WT mice were injected daily intraperitoneally with either vehicle (saline/10% DMSO), 1 mg per kg body weight CGS21680, or 1 mg per kg body weight CL 316,243 for 10 days.

CD73^{-/-} mice were provided by J. Schrader.

Pharmacological activation of energy expenditure. Eight-week-old male C57Bl/6 WT mice were injected subcutaneously five minutes before measurement. Propranolol (5 mg per kg body weight) was given 20 min before noradrenaline (1 mg per kg body weight) or CGS21680 (1 mg per kg body weight). Oxygen consumption was measured for 100 s every 5 min for 3 h with Phenomaster (TSE Systems). Time-course (Fig. 2d, g; Extended Data Fig. 6d) and relative increase (dO₂, relative to respective *t* = 0) (Fig. 2b, e, h and 4d; Extended Data Fig. 6a, e) are shown.

Physiological activation of energy expenditure. Eight-week-old male C57Bl/6 WT mice were injected subcutaneously with MSX-3 (1 mg per kg body weight) and directly put into Phenomaster cages at 4 °C. Propranolol (5 mg per kg body weight) was given 20 min before vehicle or MSX-3 injection. Animals were measured for 100 s every 5 min for 1 h. Four animals per group were analysed for both experiments.

PET/MRI of BAT activation. PET/MRI (nanoScan, Mediso Medical Imaging Systems, Hungary) studies were performed on three 8-week-old male anaesthetized C57Bl/6 WT mice (Janvier Labs, France). Subcutaneous injection of vehicle, noradrenaline (1 mg per kg body weight) or PSB-0777 (1 mg per kg body weight) was performed one minute before intraperitoneal injection of 14.7 ± 0.4 MBq of [¹⁸F]FDG (supplier: Department of Nuclear Medicine, University of Leipzig, Germany). The activity in the interscapular BAT region at 75 min post injection was expressed as mean standardized uptake value.

Glucose tolerance test. Animals were fasted for 5 h. Eight μl per g body weight of a glucose solution (2.5 g ml⁻¹) were injected intraperitoneally and glucose was measured at indicated time points post injection. Tail vein was punctured and blood was analysed with AccuChek (Aviva Nano) analyser and dipsticks (Roche).

Thermography. Thermographic images were taken from newborn littermates at room temperature with an infrared camera (IC060, Trotec) and analysed with IC-Report software 1.2 (Trotec) as previously described²³.

All studies were approved by the Landesamt für Natur, Umwelt und Verbraucherschutz, Nordrhein-Westfalen, Germany.

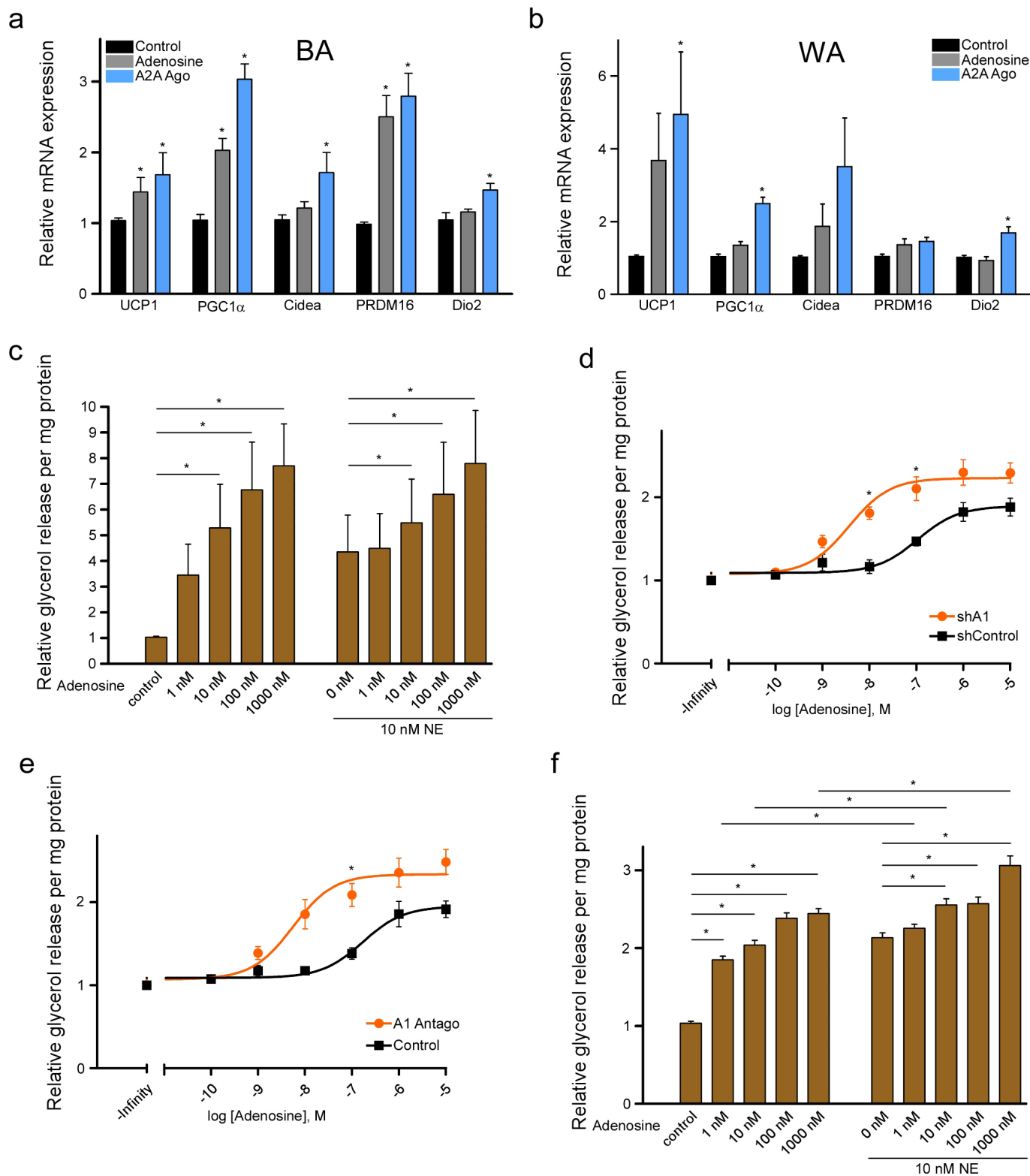
Immunohistochemistry. Five-micrometer paraffin-embedded BAT and WAT sections were blocked with 2.5% normal goat serum-PBST (phosphate-buffered saline + 0.1% Tween-20) for 1 h at room temperature. Primary antibody (UCP1, 1:50; Sigma) was applied overnight at 4 °C. After washing three times with PBST, secondary antibody against rabbit (SignalStain Boost IHC, Cell Signaling) was applied for 1 h at room temperature and developed with DAB Kit (Vector Laboratories) according to the manufacturer's instructions. Standard haematoxylin and eosin staining was performed on 5-μm paraffin-embedded BAT and WAT sections. Scale bars are 50 μm.

Statistics. To determine the group size necessary for sufficient statistical power, power analysis was performed using preliminary data.

Two-tailed student's *t*-tests were used for single comparisons and analysis of variance (ANOVA) with Bonferroni post-hoc tests for multiple comparisons. PET/MRI analysis was analysed using paired one-tailed *t*-test. *P* values below 0.05 were

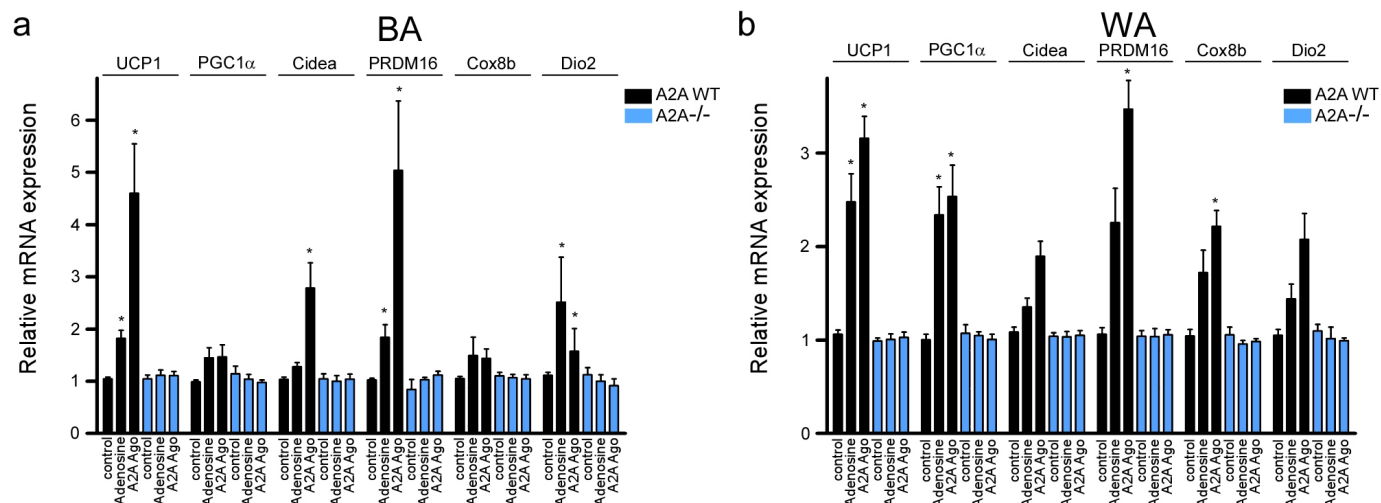
considered significant. Statistical analysis was performed with GraphPad prism 5 software. All data are represented as mean \pm s.e.m.

23. Chen, Y. *et al.* miR-155 regulates differentiation of brown and beige adipocytes via a bistable circuit. *Nature Commun.* **4**, 1769 (2013).
24. von Kügelgen, I., Allgaier, C., Schobert, A. & Starke, K. Co-release of noradrenaline and ATP from cultured sympathetic neurons. *Neuroscience* **61**, 199–202 (1994).
25. Betz, M. J. *et al.* Presence of brown adipocytes in retroperitoneal fat from patients with benign adrenal tumors: relationship with outdoor temperature. *J. Clin. Endocrinol. Metab.* **98**, 4097–4104 (2013).
26. Helenius, M., Jalkanen, S. & Yegutkin, G. Enzyme-coupled assays for simultaneous detection of nanomolar ATP, ADP, AMP, adenosine, inosine and pyrophosphate concentrations in extracellular fluids. *Biochim. Biophys. Acta* **1823**, 1967–1975 (2012).
27. Bonner, F. *et al.* Ecto-5'-nucleotidase on immune cells protects from adverse cardiac remodeling. *Circ. Res.* **113**, 301–312 (2013).



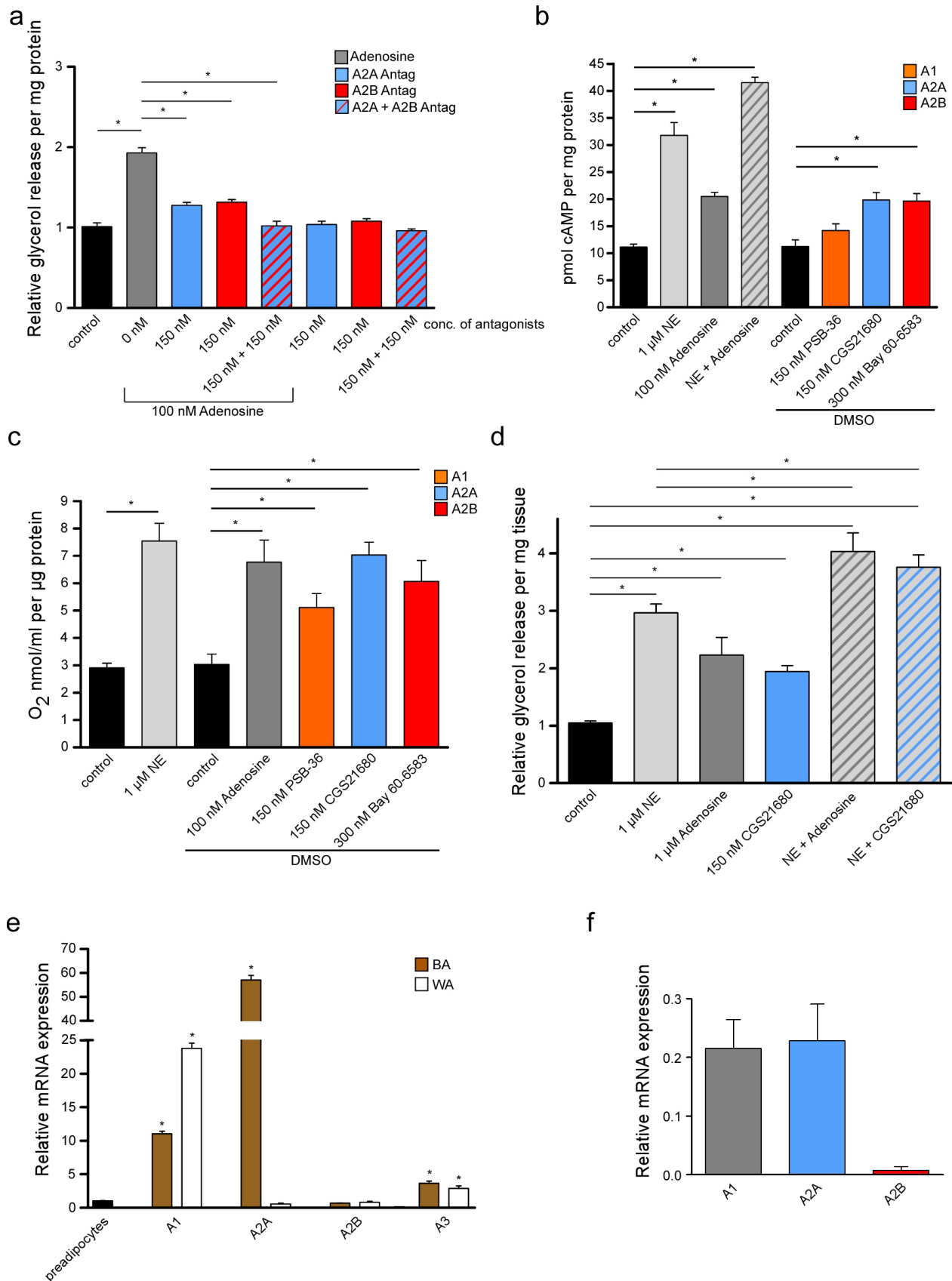
Extended Data Figure 1 | Effects of adenosine receptors in different species. **a, b**, Relative mRNA expression was analysed in (a) human brown adipocytes (BA) and (b) human white adipocytes (WA) after 8 h treatment with adenosine (BA: 70 nM; WA: 1,200 nM) for 8 h. **c**, Lipolysis in primary human brown adipocytes treated with adenosine and/or noradrenaline (NE).

d, e, Glycerol release in response to increasing concentrations of adenosine in primary murine white adipocytes after (d) knockdown of the adenosine A₁ receptor with shRNA or (e) inhibition of A₁ receptor with PSB-36 (150 nM) **f**, Adenosine-induced lipolysis in murine brown adipocytes in the presence or absence of noradrenaline. $n = 4$; * $P < 0.05$. Error bars, s.e.m.



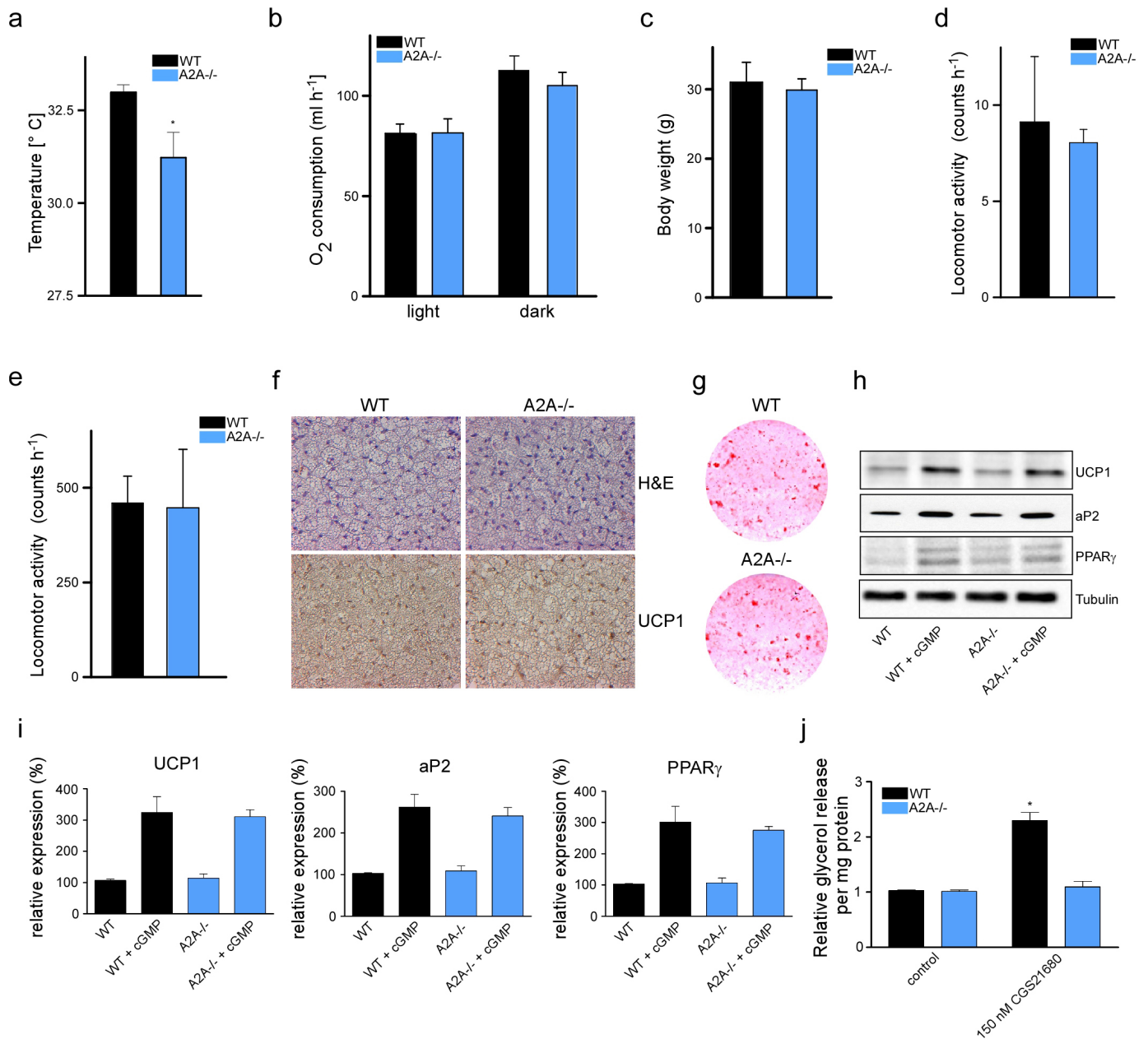
Extended Data Figure 2 | Thermogenic marker expression in A2A $^{-/-}$ and WT brown adipocytes. a, b, Murine brown adipocytes (a) or white adipocytes (b) were treated with adenosine (BA: 1 nM; WA: 100 nM) or the A_{2A}

agonist CGS21680 (150 nM) for 8 h and gene expression of indicated thermogenic markers was analysed. $n = 5$; $*P < 0.05$. Error bars, s.e.m.



Extended Data Figure 3 | Expression and function of adenosine receptors. **a**, Lipolysis in murine brown adipocytes treated with either adenosine alone or after pre-treatment with A_{2A} (MSX-2) and A_{2B} (PSB-603) receptor antagonists. **b**, cAMP levels (**b**) and oxygen consumption (**c**) of murine brown adipocytes treated with either noradrenaline, adenosine or receptor antagonists

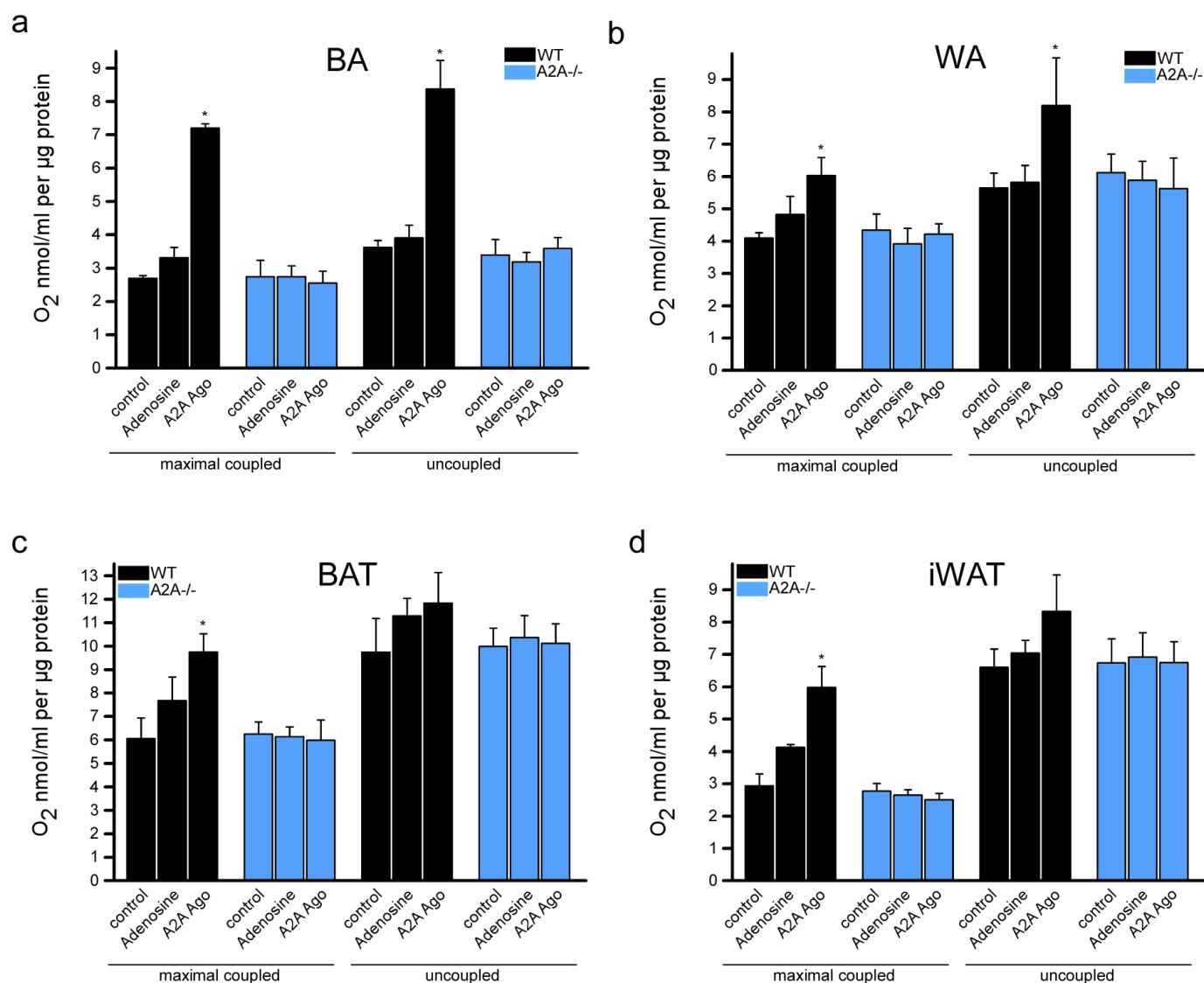
and agonists. **d**, Lipolysis of BAT explants treated with noradrenaline, adenosine or A_{2A} agonist (CGS21680). **e**, Expression of adenosine receptors in mature murine brown adipocytes and white adipocytes compared to respective pre-adipocytes. **f**, Adenosine receptor expression in BAT of Syrian golden hamster. $n = 3$ (**a–e**); $n = 8$ (**f**); $*P < 0.05$. Error bars, s.e.m.



Extended Data Figure 4 | Differentiation of A2A^{-/-} brown adipocytes.

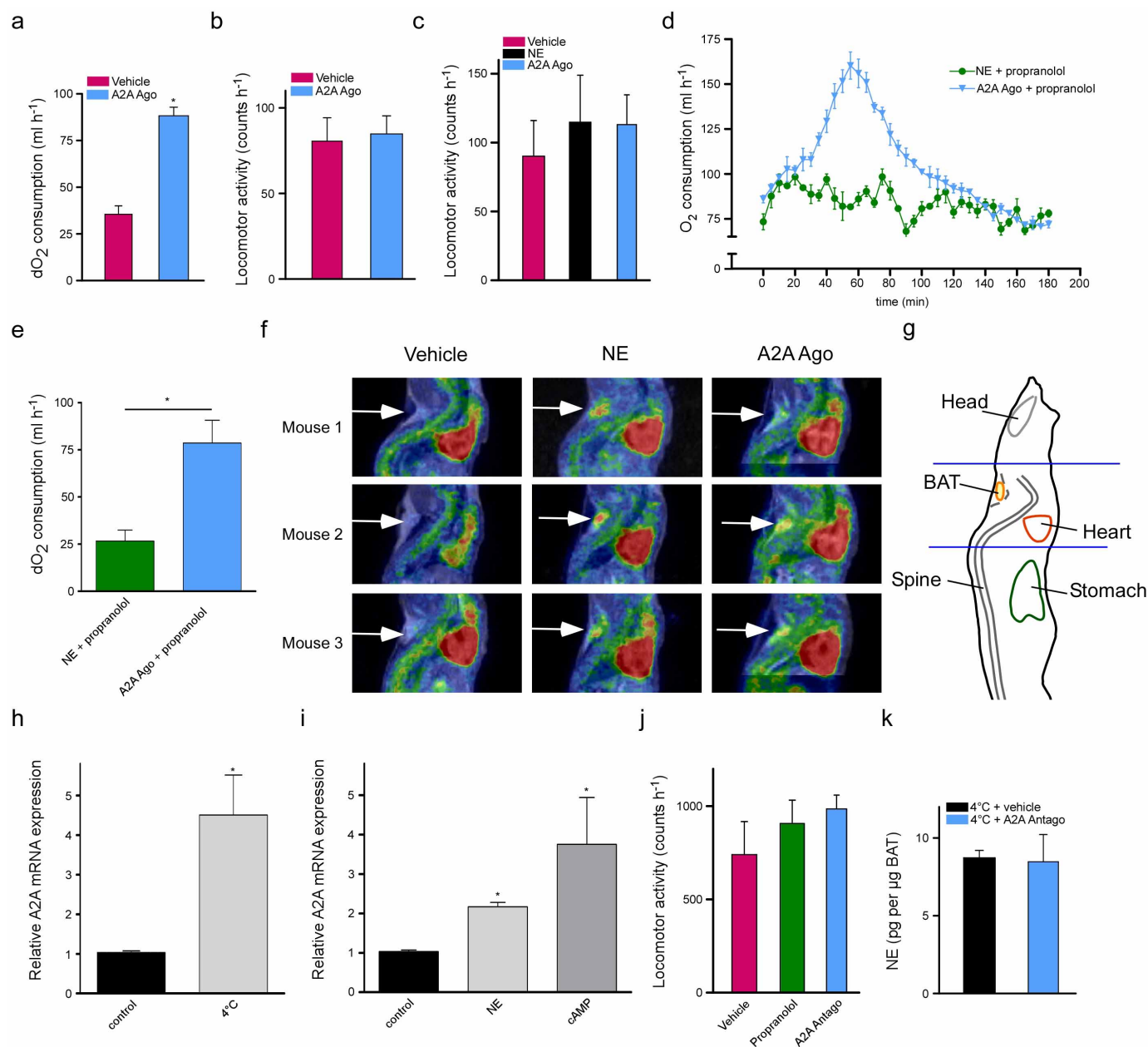
a, Quantification of neck area surface temperatures of newborn WT and A2A^{-/-} littermates. **b**, Oxygen consumption in control and A2A^{-/-} mice at thermoneutrality (30 $^{\circ}\text{C}$). **c**, Body weight of A2A^{-/-} and WT mice. **d**, **e**, Locomotor activity of mice analysed at 30 $^{\circ}\text{C}$ (**d**) or 4 $^{\circ}\text{C}$ (**e**). **f**, Representative immunohistochemistry of BAT of WT and A2A^{-/-} littermates stained with either haematoxylin and eosin or antibody against

UCP1. **g**, Representative Oil Red-O staining of differentiated WT or A2A^{-/-} brown adipocytes. **h**, Representative immunoblots of adipogenic (PPAR γ , aP2) and thermogenic marker (UCP1) expression in WT and A2A^{-/-} brown adipocytes. cGMP, 100 μM cGMP. **i**, Quantification of UCP1, aP2 and PPAR γ protein levels. **j**, Lipolysis after treatment of WT or A2A^{-/-} brown adipocytes with CGS21680. $n = 3$ (**a-e**, **i**, **j**); * $P < 0.05$. Error bars, s.e.m.



Extended Data Figure 5 | Respiration of A2A^{-/-} and WT cells and tissues.
a, b, Murine brown adipocytes (BA) (**a**) or white adipocytes (WA) (**b**) or freshly isolated BAT (**c**) and iWAT (**d**) from WT or A2A^{-/-} animals were

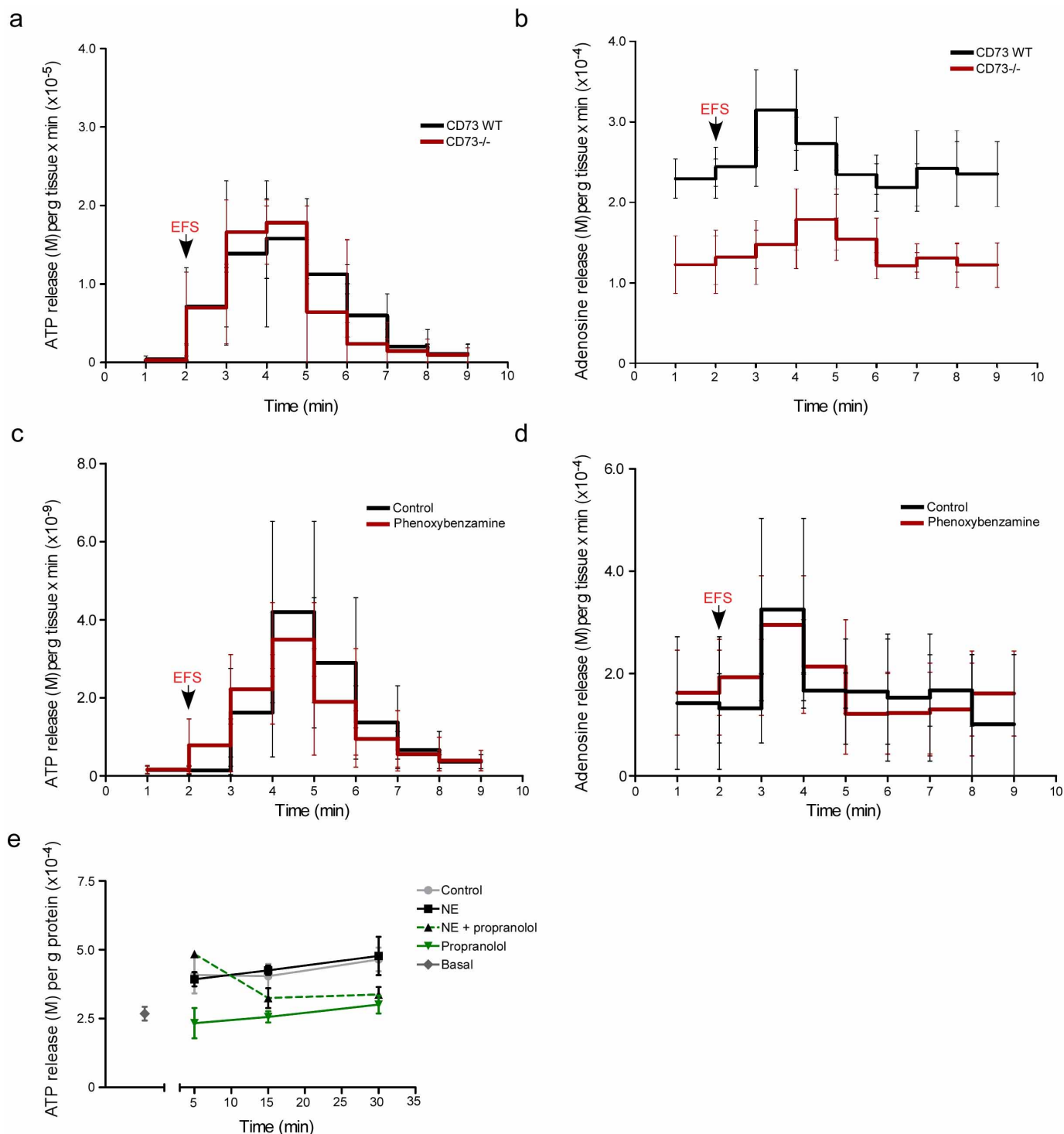
treated with adenosine (brown adipocytes and BAT: 1 nM; white adipocytes and iWAT: 100 nM) or the A_{2A} agonist CGS21680 (150 nM) and respiration was measured. $n = 4$; * $P < 0.05$. Error bars, s.e.m.



Extended Data Figure 6 | Effect of A_{2A} agonist on energy expenditure.

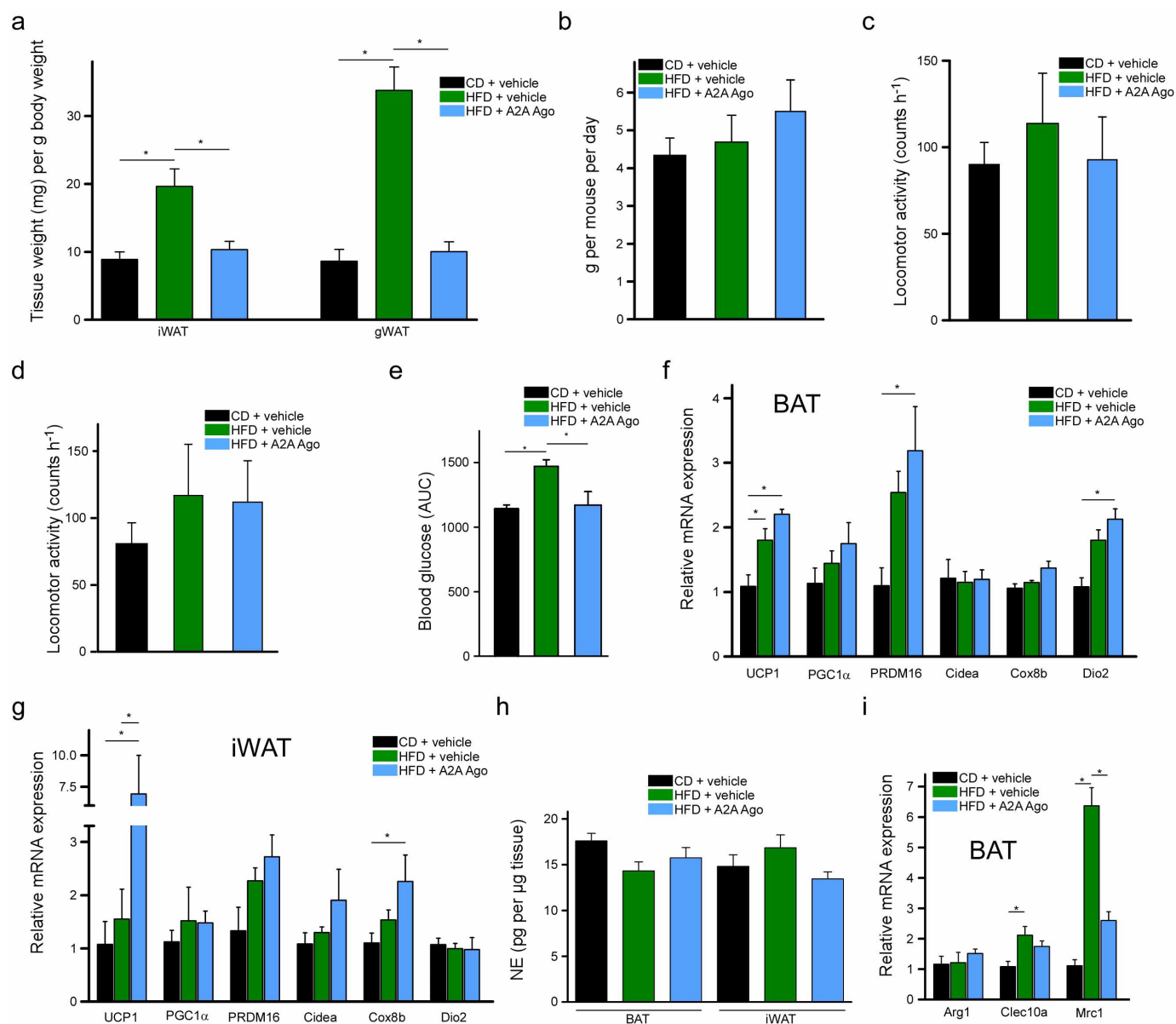
a, b, Oxygen consumption (**a**) and locomotor activity (**b**) in mice treated with A_{2A} agonist (PSB-0777). **c**, locomotor activity of mice shown in Fig. 2d, e. **d, e**, Whole body oxygen consumption in mice treated with noradrenaline or CGS21680 with propranolol pre-treatment. **f**, PET/MRI imaging of mice that were treated with vehicle, noradrenaline and PSB-0777 before injection of FDG. Arrows indicate interscapular area demonstrating increased uptake of

radioactivity by BAT. **g**, Localization of the area shown in (**f**). **h, i**, Expression of A_{2A} in BAT of mice exposed to 4 °C or room temperature (control) (**h**) and brown adipocytes after incubation with noradrenaline (1 µM) or cAMP (200 µM) (**i**). **j**, Locomotor activity of mice shown in Fig. 2g, h. **k**, Abundance of noradrenaline in mice exposed to 4 °C treated with or without A_{2A} antagonist MSX-3. *n* = 3 (**a–f, h–k**); **P* < 0.05. Error bars, s.e.m.



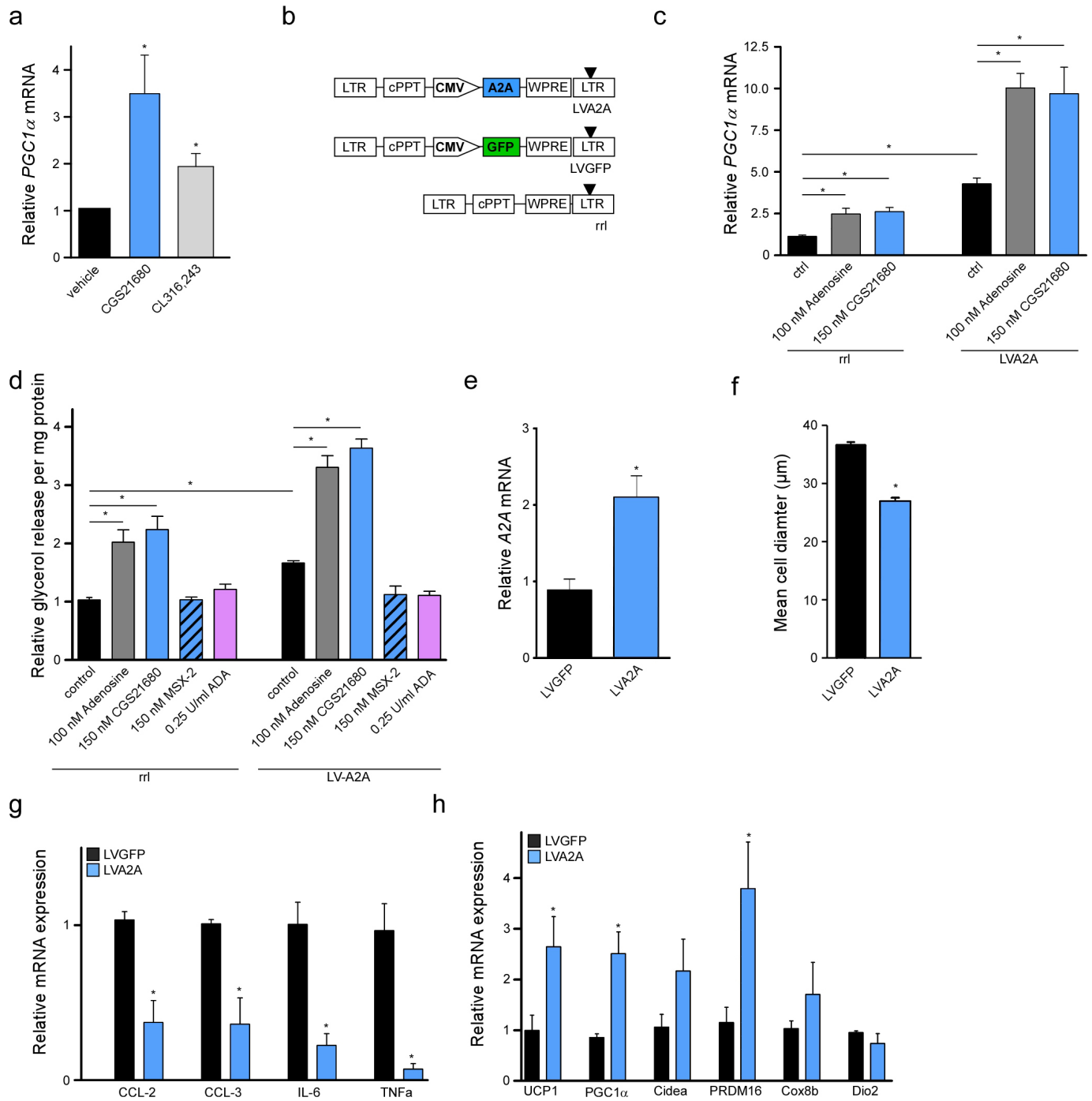
Extended Data Figure 7 | ATP and adenosine levels in BAT and brown adipocytes after stimulation. **a, b**, Concentration of ATP (**a**) and adenosine (**b**) in BAT isolated from CD73^{-/-} and WT mice subjected to EFS. **c, d**, Concentration of ATP (**c**) and adenosine (**d**) in BAT treated with

phenoxybenzamine (10 μ M) or vehicle. **e**, ATP concentrations in supernatant of murine brown adipocytes after treatment with noradrenaline in presence or absence of propranolol. $n = 3$; * $P < 0.05$. Error bars, s.e.m.



Extended Data Figure 8 | A_{2A} agonist protects mice from diet-induced obesity. **a**, Weight of iWAT and gWAT. **b**, **c**, Food intake (**b**) and locomotor activity (**c**) of mice on HFD depicted in Fig. 4a–c. **d**, Locomotor activity of mice shown in Fig. 4d. **e**, Area under the curve of glucose tolerance.

f, **g**, Thermogenic marker in BAT (**f**) or iWAT (**g**). **h**, Abundance of noradrenaline in BAT and iWAT. **i**, Alternatively activated macrophage markers in BAT. $n = 6$; $*P < 0.05$. Error bars, s.e.m.



Extended Data Figure 9 | In vivo effects of A_{2A} agonists or lentiviral A_{2A} expression. **a**, PGC1 α expression in iWAT of mice treated with A_{2A} agonist (CGS21680) or CL316,243 for 10 days. **b**, Schematic representation of lentiviral constructs. **c**, **d**, PGC1 α expression (**c**) in murine white adipocytes infected with either control virus (rrl) or LVA2A treated with adenosine or

CGS21680 and lipolysis (**d**) after treatment with adenosine, CGS21680, MSX-2 (A_{2A} antagonist) or adenosine deaminase (ADA). **e–h**, Expression of A_{2A} (**e**), adipocyte diameter (**f**), proinflammatory cytokines (**g**) and thermogenic marker genes (**h**) six weeks after injection of LVGFP or LVA2A into iWAT. $n = 4$ (**a**), $n = 3$ (**c–h**); * $P < 0.05$. Error bars, s.e.m.

Modelling human development and disease in pluripotent stem-cell-derived gastric organoids

Kyle W. McCracken¹, Emily M. Católica¹, Calyn M. Crawford¹, Katie L. Sinagoga¹, Michael Schumacher², Briana E. Rockich³, Yu-Hwai Tsai⁴, Christopher N. Mayhew¹, Jason R. Spence^{3,4}, Yana Zavros² & James M. Wells^{1,5}

Gastric diseases, including peptic ulcer disease and gastric cancer, affect 10% of the world's population and are largely due to chronic *Helicobacter pylori* infection^{1–3}. Species differences in embryonic development and architecture of the adult stomach make animal models suboptimal for studying human stomach organogenesis and pathogenesis⁴, and there is no experimental model of normal human gastric mucosa. Here we report the *de novo* generation of three-dimensional human gastric tissue *in vitro* through the directed differentiation of human pluripotent stem cells. We show that temporal manipulation of the FGF, WNT, BMP, retinoic acid and EGF signalling pathways and three-dimensional growth are sufficient to generate human gastric organoids (hGOs). Developing hGOs progressed through molecular and morphogenetic stages that were nearly identical to the developing antrum of the mouse stomach. Organoids formed primitive gastric gland- and pit-like domains, proliferative zones containing LGR5-expressing cells, surface and antral mucous cells, and a diversity of gastric endocrine cells. We used hGO cultures to identify novel signalling mechanisms that regulate early endoderm patterning and gastric endocrine cell differentiation upstream of the transcription factor NEUROG3. Using hGOs to model pathogenesis of human disease, we found that *H. pylori* infection resulted in rapid association of the virulence factor CagA with the c-Met receptor, activation of signalling and induction of epithelial proliferation. Together, these studies describe a new and robust *in vitro* system for elucidating the mechanisms underlying human stomach development and disease.

The human stomach contains a complex, three-dimensional glandular epithelium that is organized into two distinct functional domains⁵: the fundus (corpus), which is the major source of peptidases and acid, and the antrum (pylorus) that comprises a concentration of mucus-secreting cells and hormone-producing endocrine cells. Unlike other endoderm organs, little is known about signalling pathways regulating gastric development and, to our knowledge, no one has yet generated gastric tissues from human pluripotent stem cells^{6–8}. Therefore, to direct differentiation of stem cells into complex, three-dimensional gastric tissue, we needed to identify the signalling pathways that regulate several critical stages of early stomach development including: (1) posterior foregut specification and formation of the anterior gut tube; (2) gastric specification and patterning into the fundus or antrum; and (3) epithelial growth, morphogenesis, and differentiation into gastric cell lineages.

We first differentiated pluripotent stem cells into definitive endoderm⁹, which *in vivo* is then patterned along the anterior-to-posterior axis and transformed into a gut tube consisting of Sox2⁺ foregut in the anterior and Cdx2⁺ mid-hindgut in the posterior as shown in control mouse embryos (Fig. 1a). We previously demonstrated that WNT3A and FGF4 synergize to induce the morphogenesis of gut tube-like structures expressing the posterior marker CDX2 (refs 6,10). To generate foregut, from which the stomach derives, we aimed to stimulate gut tube morphogenesis

with WNT and FGF while inhibiting their ability to promote posterior fate. We found that WNT/FGF require BMP activity to initiate posterior gene expression, consistent with the known role of BMP as a posterior-defining factor^{11–13}. Specifically, inhibiting BMP signalling with the antagonist noggin (NOG) resulted in repression of the posterior marker CDX2, activation of the foregut marker SOX2 and assembly of three-dimensional foregut spheroids (Fig. 1b–d and Extended Data Fig. 1). Foregut spheroid morphogenesis was a robust process using both human embryonic stem (ES) cell and human induced pluripotent stem (iPS) cell lines (Fig. 1c, d and Extended Data Fig. 2). Thus, we identified a new epistatic relationship between WNT, FGF and BMP in which all three pathways cooperate to promote a mid-hindgut fate, but WNT and FGF act separately from BMP to drive morphogenesis of gut tube structures.

The subsequent events of stomach development *in vivo* are posterior patterning of the foregut and specification of the fundic and antral domains of the stomach. To direct spheroids into a posterior foregut fate (indicated in mouse embryos by co-expression of Sox2 and Hnf1 β ; Fig. 1e), we focused on retinoic acid signalling given its role in development of posterior-foregut-derived organs^{14–16}. Exposing definitive endoderm to retinoic acid for 24 h on the final day (day 5–6) of the patterning/spheroid generation stage resulted in the formation of SOX2/HNF1 β ⁺ posterior foregut spheroids (Fig. 1f, g and Extended Data Fig. 3). In the mouse embryo, the posterior foregut undergoes morphogenesis and is subdivided into the Sox2⁺/Pdx1[–] fundus, Sox2/Pdx1⁺ antrum, Pdx1/Ptf1a⁺ pancreas, and Pdx1/Cdx2⁺ duodenum (Fig. 2b). To promote three-dimensional growth and morphogenesis, we transferred posterior foregut spheroids to a semisolid matrix and found that an additional 72 h of retinoic acid (day 6–9) caused a >100-fold increase in *PDX1* messenger RNA levels while maintaining high SOX2 expression (Fig. 2c, d), indicating specification into antrum. Importantly, the retinoic acid treatment did not promote a pancreatic fate⁸, since expression of the pancreas-specific marker *PTF1A* (ref. 17) was not induced (Fig. 2d). These data demonstrate that the combination of retinoic acid signalling with three-dimensional growth efficiently direct posterior foregut spheroids into a SOX2/PDX1⁺ epithelium indicative of a gastric antrum fate.

The next stages of stomach development are characterized by growth of a pseudo-stratified epithelium into an elaborate glandular epithelium. We explored various growth conditions and found that high concentrations of EGF (100 ng ml^{–1}) were sufficient to promote robust outgrowth of SOX2/PDX1⁺ spheroids into hGOs. Over the course of 3–4 weeks, spheroids (<100 μ m in diameter) underwent marked expansion into organoids (2–4 mm in diameter) containing a complex columnar epithelium (Fig. 2e). EGF was required for the initial outgrowth of foregut spheroids as well as their expansion and morphogenesis at later stages (Extended Data Fig. 4c), revealing a new role for EGF receptor signalling during embryonic stomach formation that is temporally distinct from its postnatal function as a trophic factor^{18,19}. hGO development

¹Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, Ohio 45229-3039, USA. ²Department of Molecular and Cellular Physiology, University of Cincinnati, Cincinnati, Ohio 45267, USA. ³Department of Cell and Developmental Biology, University of Michigan Medical School, Ann Arbor, Michigan 48109-2200, USA. ⁴Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan 48109-2200, USA. ⁵Division of Endocrinology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, Ohio 45229-3039, USA.

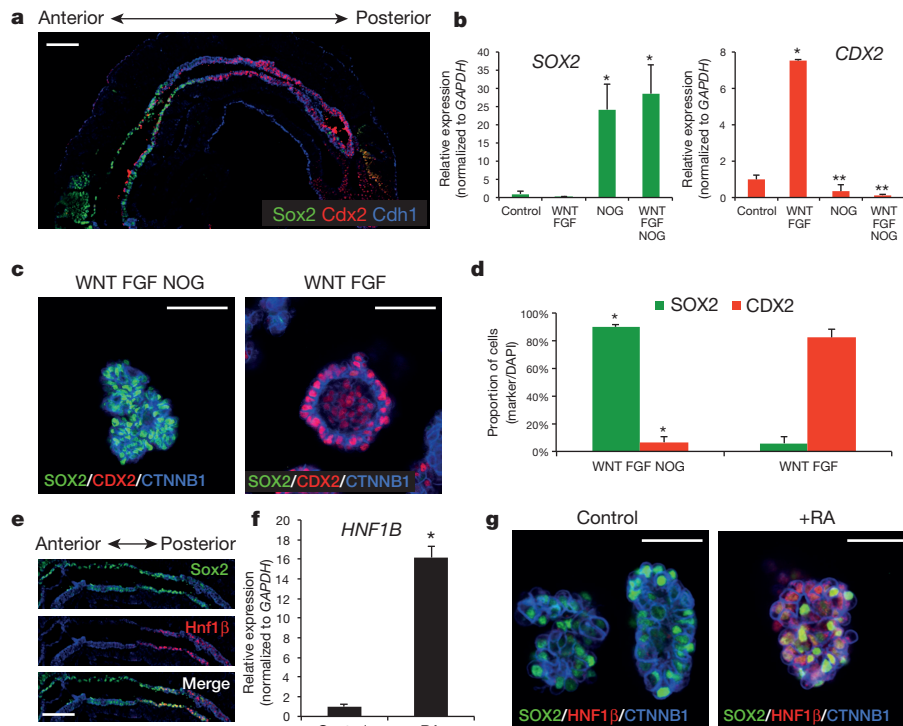


Figure 1 | Generation of three-dimensional posterior foregut spheroids.

a, Sox2 marks foregut endoderm and Cdx2 marks mid/hindgut endoderm in E8.5 (14-somite stage) mouse embryo. **b, c**, Quantitative PCR (qPCR) analysis (**b**) and wholemount immunostaining (**c**) for patterning markers at day 6 in human pluripotent stem-cell definitive endoderm cultures exposed to 3 days in media alone (control) or with the indicated growth factors/antagonists. WNT3A (WNT) and FGF4 (FGF) induced CDX2 expression whereas the BMP antagonist NOG repressed CDX2 and induced high levels of the foregut marker SOX2. Results are normalized to expression in control endoderm (stage-matched, no-growth-factor-treated). * $P < 0.05$ compared to control; ** $P < 0.005$ compared to WNT/FGF; two-tailed Student's *t*-test; $n = 3$ biological replicates per condition, data representative of 6 independent experiments. **d**, Quantitation of SOX2- and CDX2-expressing cells in day 6

spheroids generated in hindgut (WNT/FGF4) and foregut (WNT/FGF4/NOG) patterning conditions. Data are expressed as the percentage of cells expressing indicated markers, normalized to the total number of cells in the spheroids.

* $P < 1.0 \times 10^{-6}$; two-tailed Student's *t*-test; $n = 5$ biological replicates per condition, data representative of 3 independent experiments. **e**, The posterior foregut in the E8.5 mouse embryo expressed both Sox2 and Hnf1 β .

f, g, Exposing cultures to retinoic acid (RA) on the final day (day 5–6) of the spheroid generation step induced expression of HNF1 β in SOX2-expressing epithelium, measured by both qPCR (**f**) and wholemount immunofluorescent staining (**g**) at day 6, indicating the formation of posterior foregut spheroids.

* $P < 0.005$; two-tailed Student's *t*-test; $n = 3$ biological replicates per condition, data representative of 3 independent experiments. Scale bars, 100 μ m (**a** and **e**) and 50 μ m (**c** and **g**). Error bars represent s.d.

was remarkably similar to *in vivo* stomach organogenesis. At early stages (embryonic day (E)12–14 in mouse and day-13 hGOs), both epithelia were pseudo-stratified, contained mitotic nuclei concentrated towards the apical surface indicating interkinetic nuclear migration, and were appropriately polarized and contained deep elaborations of atypical protein kinase C (aPKC⁺) apical membrane (Extended Data Fig. 4b)²⁰. At later stages (E16.5–postnatal day (P)12 in mouse, day-13–34 hGOs), the *in vivo* antrum transformed into a simple columnar epithelium exhibiting a highly structured organization, and the hGOs underwent similar folding and formed immature pit and gland domains (Fig. 2e, f and Extended Data Fig. 4a).

Molecular markers that define the developing mouse antrum *in vivo* showed analogous temporal and spatial expression patterns in developing hGOs. At early stages (E12–14 in mouse and 13-day hGOs) the transcription factors SOX2, PDX1, GATA4 and KLF5 were all co-expressed in the immature, pseudo-stratified epithelium of the antrum (Extended Data Fig. 4). However, at later stages (E18–P12 in mouse and day-34 hGOs), SOX2 was downregulated as the epithelium formed glands and pits, whereas the expression of the other factors was maintained. On the basis of these data, the day-13 hGOs represent a developmental stage similar to the E12–14 mouse antrum, whereas day-34 hGOs are more comparable to the late-fetal/early-postnatal antrum. Furthermore, the early spheroid mesoderm expanded during organoid differentiation, expressed mesenchymal transcription factors including FOXF1 and BAPX1 (ref. 21), and differentiated into VIM⁺ submucosal fibroblasts and a smaller number of ACTA2⁺ subepithelial myofibroblasts

(Extended Data Figs 4 and 5). Thus, we conclude that hGOs recapitulate normal embryonic development, and that the molecular and morphogenetic processes that occur during antrum development are conserved between rodents and humans.

RNA-sequencing (RNA-seq) analysis showed that day-34 organoids and human fetal stomach tissue share a very similar transcriptional profile, which is distinct from human fetal intestine (Extended Data Fig. 6b). The antral region in the embryonic and postnatal stomach can be distinguished from the fundus by its expression of PDX1 (Fig. 2f), as well as by a concentration of certain cell types. Antral cell types include mucous cells, which secrete the protective mucus lining, and several endocrine lineages that regulate gastrointestinal physiology and metabolic homeostasis (Fig. 3a). By day 27, hGOs robustly expressed numerous transcripts that mark differentiated antral cell types (Extended Data Fig. 6a), including surface mucous cells (*MUC5AC*, *TFF1/3* and *GKN1*) and antral gland cells (*TFF2*), but not cell types associated with the fundus/corpus such as parietal cells (*ATP4A/B*) and chief cells (*MIST1* (also known as *BHLHA15*)), or intestinal goblet cells (*MUC2*). At day 34, histological analysis revealed that both mucous cell lineages were abundant in the hGOs (Fig. 3b–f): surface mucous cells with tall columnar morphology and apical *MUC5AC* expression; and *MUC6*⁺ antral gland cells that were concentrated towards the base of the glands. Parietal cells (*ATP4B*) were undetectable in the hGOs, and the lack of fundus cell types coupled with ubiquitous PDX1 expression are consistent with an antral identity (Fig. 2 and Extended Data Fig. 4). The adult antrum contains pure glands with only antral lineages and mixed glands containing

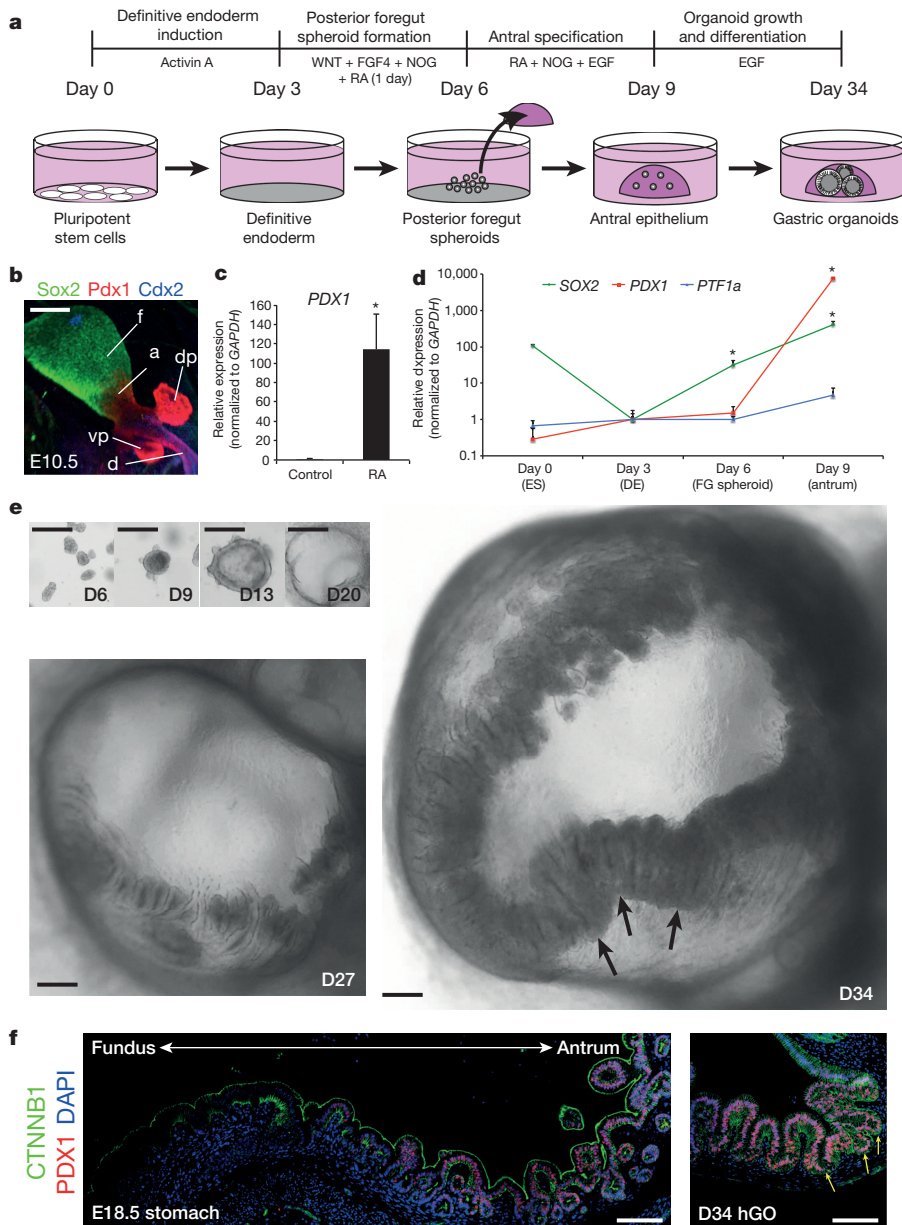


Figure 2 | Specification and growth of human antral gastric organoids. **a**, Schematic representation of the *in vitro* culture system used to direct the differentiation of pluripotent stem cells into three-dimensional gastric organoids. **b**, Defining molecular domains of the posterior foregut in E10.5 mouse embryos for Sox2, Pdx1 and Cdx2; Sox2/Pdx1, antrum (a); Sox2, fundus (f); Pdx1, dorsal and ventral pancreas (dp and vp); Pdx1/Cdx2, duodenum (d). **c**, Posterior foregut spheroids exposed for three days to retinoic acid (2 μ M) exhibited >100-fold induction of PDX1 compared to control spheroids, measured by qPCR at day 9. * P < 0.05; two-tailed Student's *t*-test; n = 3 biological replicates per condition, data representative of 4 independent experiments. **d**, Time course qPCR analysis of antral differentiation (according to protocol detailed in Fig. 2a) demonstrated sequential activation of SOX2 at day 6 (posterior foregut (FG) endoderm), followed by induction of PDX1 at day 9 (presumptive antrum). Day-9 antral spheroids had a 500-fold increase in SOX2 and a 10,000-fold increase in PDX1 relative to day-3 definitive endoderm (DE). * P < 0.05; two-tailed Student's *t*-test; n = 3 biological replicates per time point, data representative of 2 independent experiments. The pancreatic marker PTF1A was not significantly increased. **e**, Stereomicrographs showing morphological changes during growth of gastric organoids. By 4 weeks, the epithelium of hGOs exhibited a complex folded and glandular architecture (arrows). **f**, Comparison of mouse stomach at E18.5 and day-34 hGOs. Pdx1 was highly expressed in the mouse antrum but excluded from the fundus. Human gastric organoids expressed PDX1 throughout the epithelium and exhibited morphology similar to the late gestational mouse antrum (arrows). Scale bars, 100 μ m (b and f) and 250 μ m (e). Error bars represent s.d.

both antral and fundic lineages²². However, PDX1 expression remains restricted to antral cell types even in mixed glands²², suggesting the possibility that PDX1⁺ progenitor cells only give rise to antrum-specific cell types postnatally. Thus, we conclude that the PDX1-expressing hGOs are a pure antrum/pylorus cell population.

The antrum also contains LGR5-expressing stem cells²³ and specialized endocrine cells. At day 34, hGOs contained a SOX9⁺ proliferative progenitor zone as well as LGR5⁺ cells at the base of the glands (Fig. 3d–f and Fig. 4a), identified using a transgenic *LGR5-eGFP* reporter (containing enhanced green fluorescent protein) stem-cell line (Extended Data Fig. 7). hGOs contained a complete spectrum of endocrine cells normally found in the antrum that were positive for the hormones gastrin, ghrelin, somatostatin and serotonin (Fig. 3g). Notably, we observed that high levels of EGF repressed expression of the proendocrine transcription factor *NEUROG3* (refs 24,25) and the formation of endocrine cells, whereas low concentrations of EGF (10 ng ml^{−1}) supported endocrine cell formation (Extended Data Fig. 8b–d). Transient overexpression of *NEUROG3* was sufficient to overcome the endocrine inhibitory effect of EGF, resulting in robust differentiation of gastric endocrine cells (Extended Data Fig. 8e–g). In summary, hGOs have a high degree of cellular complexity that rivals that of the human antrum, they can be

experimentally manipulated, and they therefore represent the first *in vitro* model of human stomach development.

Clinical evidence indicates that predominant colonization of the antrum has an important role in *H. pylori*-mediated disease^{26,27}. Thus, we tested whether hGOs could be used to model the pathophysiological response of the human stomach to *H. pylori*. To mimic the normal host–pathogen interface, we microinjected *H. pylori* directly into the lumen of the epithelium and measured epithelial signalling and proliferation. Within 24 h, bacteria were observed tightly associated with the hGO epithelium (Fig. 4a and Extended Data Fig. 9b) and we observed robust epithelial responses including phosphorylation of c-Met²⁸ (Fig. 4b and Extended Data Fig. 9c) and a twofold increase in epithelial cell proliferation (Fig. 4c). The *H. pylori* virulence factor CagA plays a pivotal role in the aetiology of disease. Consistent with published studies²⁹, we observed that CagA translocated into organoid epithelial cells and formed a complex with the c-Met receptor (Fig. 4b). Furthermore, the epithelial response was lost when hGOs were injected with a strain of *H. pylori* lacking CagA, demonstrating the dependence on CagA for *H. pylori*-mediated human pathogenesis. Thus, the pathophysiological response of hGOs makes them an unprecedented model for elucidating the initiating events of human gastric disease mediated by *H. pylori*.

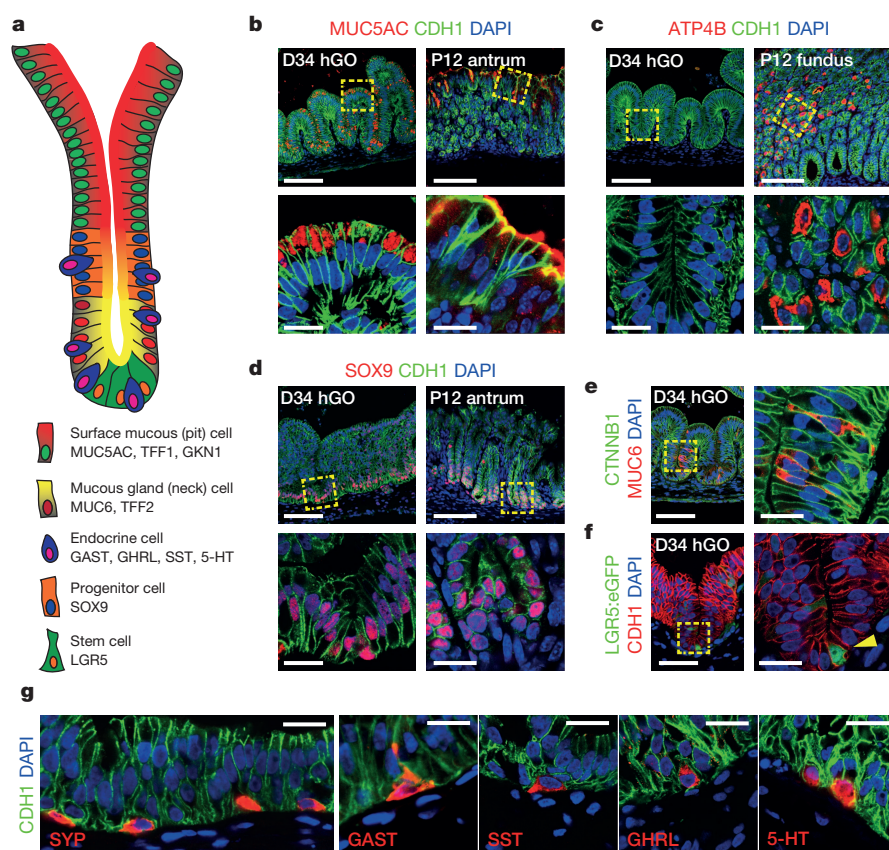


Figure 3 | hGOs contain differentiated antral cell types. **a**, Schematic representation of a typical antral gland showing normal cell types and associated molecular markers. **b–g**, Immunofluorescent staining demonstrated that day-34 hGOs consisted of normal cell types found in the antrum, but not the fundus. The hGO epithelium contained surface mucous cells that express MUC5AC (**b**, left), similar to the P12 mouse antrum (**b**, right), but not ATP4B-expressing parietal cells (**c**, left) that characterize the fundus (**c**, right). SOX9⁺ cells were found at the base of the hGO epithelium (**d**, left), similar to the progenitor cells found in the P12 antrum (**d**, right). Furthermore, hGOs contained MUC6⁺ antral gland cells (**e**) and LGR5-expressing cells (yellow arrow) (**f**). Boxed regions in **b–f** are shown as high magnification images below (**b**, **c**, **d**) or to the right (**e**, **f**) of the original. **g**, Day-34 hGOs also contained endocrine cells (SYP) that expressed the gastric hormones GAST, SST, GHRL and serotonin (5-HT). Scale bars, 100 μm (original images in **b–f**) and 20 μm (magnified images in **b–f** and **g**). Marker expression data are representative from a minimum of 10 independent experiments, except LGR5-eGFP data, which is a representative example from two separate experiments. DAPI, 4',6-diamidino-2-phenylindole.

In summary, hGOs represent the first, to our knowledge, human gastric tissue fully derived *in vitro* (summarized in Extended Data Figure 10) and are one of the most physiologically complex micro-organ systems

yet established. The hGOs undergo normal stages of *in vivo* differentiation, comprise an array of cell types that constitute the normal antral epithelium, and contain a complex three-dimensional organization. We have used hGOs as an *in vitro* system to identify signalling mechanisms that regulate human stomach development and physiology, and we have modelled the pathophysiological response of the gastric epithelium to *H. pylori*. Thus, hGOs should present new opportunities for drug discovery and modelling the early stages of gastric disease, including cancer. Moreover, this is, to our knowledge, the first three-dimensional production of a human embryonic foregut and represents a promising starting point for the three-dimensional generation of other foregut organ tissues including fundus, lungs and pancreas.

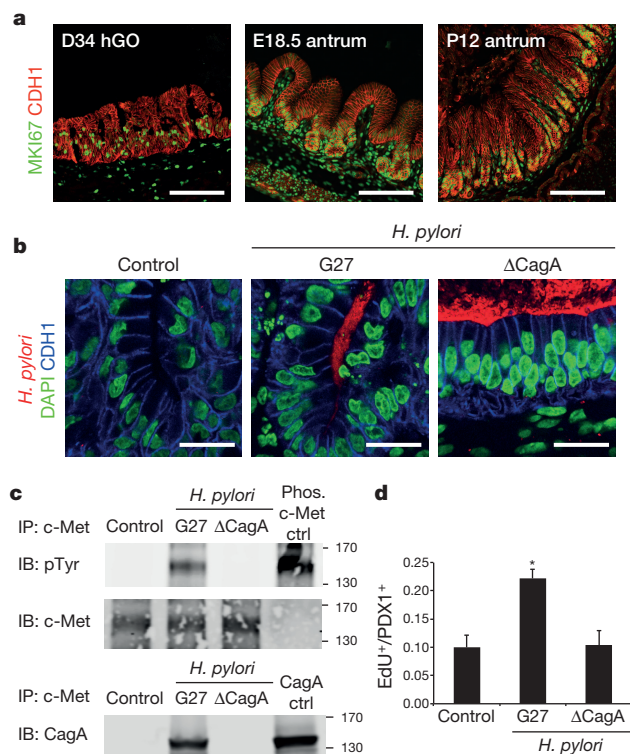


Figure 4 | hGOs exhibit acute responses to *H. pylori* infection. **a**, Day-34 hGOs contained a zone of MKI67⁺ proliferative cells similar to the embryonic (E18.5) and postnatal (P12) mouse antrum. **b**, Using hGOs to model human-specific disease processes of *H. pylori* infection. Pathogenic (G27) and attenuated (Δ CagA) bacteria were microinjected into the lumen of hGOs and after 24 h, bacteria (both G27 and Δ CagA strains) were tightly associated with the apical surface of the hGO epithelium. **c**, Immunoprecipitation (IP) for the oncogene c-Met demonstrates that *H. pylori* induced a robust activation (tyrosine phosphorylation (pTyr)) of c-Met, and this is a CagA-dependent process. Furthermore, CagA immunoprecipitated with c-Met, suggesting that these proteins interact in hGO epithelial cells. Phosphorylated c-Met (phos. c-MET) and CagA control lysates were not immunoprecipitated but used to confirm molecular masses. The molecular mass markers are indicated (130 and 170 kilodaltons (kDa)) and shown in Extended Data Fig. 9c. IB, immunoblotting. **d**, Within 24 h, *H. pylori* infection caused a CagA-dependent twofold increase in the number of proliferating cells in the hGO epithelium, measured by 5-ethynyl-2'-deoxyuridine (EdU) incorporation. **P* < 0.05; two-tailed Student's *t*-test; *n* = 3 biological replicates per condition, data representative of 4 independent experiments. Scale bars, 100 μm (**a**) and 20 μm (**b**). Error bars represent s.e.m.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 12 February; accepted 12 September 2014.

Published online 29 October 2014.

- Wen, S. & Moss, S. F. *Helicobacter pylori* virulence factors in gastric carcinogenesis. *Cancer Lett.* **282**, 1–8 (2009).
- Yuan, Y., Padol, I. T. & Hunt, R. H. Peptic ulcer disease today. *Nature Clin. Pract. Gastroenterol. Hepatol.* **3**, 80–89 (2006).
- Parkin, D. M. The global health burden of infection-associated cancers in the year 2002. *Int. J. Cancer* **118**, 3030–3044 (2006).
- Peek, R. M. *Helicobacter pylori* infection and disease: from humans to animal models. *Dis. Model. Mech.* **1**, 50–55 (2008).
- Mills, J. C. & Shivdasani, R. A. Gastric epithelial stem cells. *Gastroenterology* **140**, 412–424 (2011).
- Spence, J. R. *et al.* Directed differentiation of human pluripotent stem cells into intestinal tissue *in vitro*. *Nature* **470**, 105–109 (2011).
- Si-Tayeb, K. *et al.* Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. *Hepatology* **51**, 297–305 (2010).
- D'Amour, K. A. *et al.* Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nature Biotechnol.* **24**, 1392–1401 (2006).
- D'Amour, K. A. *et al.* Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nature Biotechnol.* **23**, 1534–1541 (2005).
- McCracken, K. W., Howell, J. C., Spence, J. R. & Wells, J. M. Generating human intestinal tissue from pluripotent stem cells *in vitro*. *Nature Protocols* **6**, 1920–1928 (2011).
- Kumar, M., Jordan, N., Melton, D. & Grapin-Botton, A. Signals from lateral plate mesoderm instruct endoderm toward a pancreatic fate. *Dev. Biol.* **259**, 109–122 (2003).
- Tiso, N., Filippi, A., Pauls, S., Bortolussi, M. & Argenton, F. BMP signalling regulates anteroposterior endoderm patterning in zebrafish. *Mech. Dev.* **118**, 29–37 (2002).
- Green, M. D. *et al.* Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. *Nature Biotechnol.* **29**, 267–272 (2011).
- Wang, Z., Dollé, P., Cardoso, W. V. & Niederreither, K. Retinoic acid regulates morphogenesis and patterning of posterior foregut derivatives. *Dev. Biol.* **297**, 433–445 (2006).
- Martín, M. *et al.* Dorsal pancreas agenesis in retinoic acid-deficient *Raldh2* mutant mice. *Dev. Biol.* **284**, 399–411 (2005).
- Molotkov, A., Molotkova, N. & Ducrest, G. Retinoic acid generated by *Raldh2* in mesoderm is required for mouse dorsal endodermal pancreas development. *Dev. Dyn.* **232**, 950–957 (2005).
- Kawaguchi, Y. *et al.* The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nature Genet.* **32**, 128–134 (2002).
- Johnson, L. R. & Guthrie, P. D. Stimulation of rat oxyntic gland mucosal growth by epidermal growth factor. *Am. J. Physiol.* **238**, G45–G49 (1980).
- Majumdar, A. P. Postnatal undernutrition: effect of epidermal growth factor on growth and function of the gastrointestinal tract in rats. *J. Pediatr. Gastroenterol. Nutr.* **3**, 618–625 (1984).
- Grosse, A. S. *et al.* Cell dynamics in fetal intestinal epithelium: implications for intestinal growth and morphogenesis. *Development* **138**, 4423–4432 (2011).
- Verzi, M. P. *et al.* Role of the homeodomain transcription factor Bapx1 in mouse distal stomach development. *Gastroenterology* **136**, 1701–1710 (2009).
- Choi, E. *et al.* Cell lineage distribution atlas of the human stomach reveals heterogeneous gland populations in the gastric antrum. *Gut* (2014).
- Barker, N. *et al.* Lgr5⁺ stem cells drive self-renewal in the stomach and build long-lived gastric units *in vitro*. *Cell Stem Cell* **6**, 25–36 (2010).
- Jenny, M. *et al.* Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *EMBO J.* **21**, 6338–6347 (2002).
- Lee, C. S., Perreault, N., Brestelli, J. E. & Kaestner, K. H. Neurogenin 3 is essential for the proper specification of gastric enteroendocrine cells and the maintenance of gastric epithelial cell identity. *Genes Dev.* **16**, 1488–1497 (2002).
- Olbe, L., Hamlet, A., Dalenbäck, J. & Fändriks, L. A mechanism by which *Helicobacter pylori* infection of the antrum contributes to the development of duodenal ulcer. *Gastroenterology* **110**, 1386–1394 (1996).
- Xia, H. H. *et al.* Antral-type mucosa in the gastric incisura, body, and fundus (antralization): a link between *Helicobacter pylori* infection and intestinal metaplasia? *Am. J. Gastroenterol.* **95**, 114–121 (2000).
- Churin, Y. *et al.* *Helicobacter pylori* CagA protein targets the c-Met receptor and enhances the mitogenic response. *J. Cell Biol.* **161**, 249–255 (2003).
- Peek, R. M. *et al.* *Helicobacter pylori* cagA⁺ strains and dissociation of gastric epithelial cell proliferation from apoptosis. *J. Natl. Cancer Inst.* **89**, 863–868 (1997).

Acknowledgements We thank A. Zorn, J. Whitsett, N. Shroyer, the Pluripotent Stem Cell Facility and members of the Wells and Zorn laboratories for reagents and feedback. We also thank M. Kofron for assistance with confocal imaging and T. Westbrook for providing the pInducer20 vector. We thank R. Peek for assistance with analysing electron micrograph images. This work was supported by National Institutes of Health grants R01DK080823, R01DK092456 and K01DK091415, NIGMS Medical Scientist Training Program T32 GM063483, and the American Gastroenterological Association: Robert and Sally Funderburg Research Award in Gastric Cancer. We also acknowledge core support from the Cincinnati Digestive Disease Center Award (P30 DK0789392), Clinical Translational Science Award (U54 RR025216), the Michigan Gastrointestinal Peptide Research Center (MGPRC: NIDDK 5P30DK034933), and technical support from CCHMC Confocal Imaging Core, CCHMC Pathology Core, and CCHMC Viral Vector Core.

Author Contributions K.W.M. and J.M.W. conceived the study and experimental design, performed and analysed experiments and co-wrote the manuscript. Y.Z. designed, performed and helped analyse *H. pylori* experiments. E.M.C., C.M.C., K.L.S. and M.S. performed experiments. C.N.M. generated and characterized the iPS cell line. B.E.R., Y.-H.T. and J.R.S. designed, generated and characterized the LGR5-eGFP reporter hES cell line and performed RNA-seq experiments and analysis. All authors contributed to the writing or editing of the manuscript.

Author Information The RNAseq data from hGOs have been deposited in ArrayExpress with accession number E-MTAB-2885. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.M.W. (james.wells@cchmc.org) or Y.Z. (zavrosya@ucmail.uc.edu).

METHODS

Pluripotent stem-cell culture. Human ES cell lines WA01 (H1) and WA09 (H9) were obtained from WiCell. ES and iPS cell lines were maintained as colonies in feeder-free conditions on human-ES-cell-qualified Matrigel (BD Biosciences) in mTesR1 media (Stem Cell Technologies). Cells were routinely passaged every 4 days with dispase (Invitrogen).

Differentiation of definitive endoderm. For differentiation, pluripotent stem cells were plated as single cells in a Matrigel-coated 24-well dish using accutase (Stem Cell Technologies), at a density of 150,000 cells per well in mTesR1 with ROCK inhibitor Y-27632 (10 μ M; Stemgent). On the next day, stem cells were differentiated to definitive endoderm as previously described^{9,30}. Cells were exposed to activin A (100 ng ml⁻¹; Cell Guidance Systems) for 3 days in RPMI 1640 media (Invitrogen) containing increasing concentrations of 0%, 0.2% and 2.0% defined FBS (dFBS; Invitrogen). In addition, BMP4 (50 ng ml⁻¹; R&D Systems) was added on the first day of definitive endoderm induction.

Endoderm patterning and foregut spheroid generation. After definitive endoderm induction, cells were cultured in RPMI 1640 media with 2.0% dFBS and the indicated combinations of growth factors and/or chemical agonist: WNT3A (500 ng ml⁻¹; R&D Systems); CHIR99021 (2 μ M; Stemgent); FGF4 (500 ng ml⁻¹; R&D Systems); and NOG (200 ng ml⁻¹; R&D Systems). The media was changed every day. After three days, the combination of WNT3A (or CHIR99021), FGF4 and NOG resulted in floating foregut spheroids in the culture wells. To induce a posterior fate in foregut endoderm, retinoic acid (2 μ M; Sigma Aldrich) was added on the third day of WNT/FGF/NOG treatment.

Three-dimensional culture of gastric organoids. The spheroids were transferred to a three-dimensional *in vitro* culture system as previously described^{6,10}. In brief, spheroids were collected, resuspended in 50 μ l Matrigel (BD Biosciences), and plated in a three-dimensional droplet. After Matrigel was allowed to solidify for 10–15 min in a tissue culture incubator, spheroids were overlaid with gut media: Advanced DMEM/F12 with N2 (Invitrogen), B27 (Invitrogen), L-glutamine, 10 μ M HEPES, penicillin/streptomycin, and EGF (100 ng ml⁻¹; R&D Systems). For the first 3 days, retinoic acid and NOG were added to the gut media. Media was replaced every 3–4 days, as necessary. At day 20, organoids were collected and re-plated in fresh Matrigel at a dilution of ~1:12.

Generation of doxycycline-inducible *NEUROG3* stem-cell line. To generate the overexpression construct, human *NEUROG3* complementary DNA (Dana-Farber/Harvard Cancer Center DNA Resource Core; clone HsCD00345898) was cloned into pInducer20 lentiviral vector (provided by T. Westbrook³¹) using Gateway Cloning (Invitrogen) methods. High-titre lentiviral particles were produced by the CCHMC Viral Vector Core. H1 ES cells were dissociated with Accutase, plated as a single-cell suspension in mTesR1 with 10 μ M Y-27632, and exposed to lentivirus for 4 hours. mTesR1 was replaced daily and after 2 days, G418 (200 μ g ml⁻¹) was added to the media to select for integrated clones. G418-resistant cells were maintained in antibiotic indefinitely, but were otherwise cultured and passaged normally.

Generation of *LGR5-eGFP* BAC transgenic reporter ES cell line. Bacterial artificial chromosome (BAC) RP11-59F15 was obtained from the Children's Hospital Oakland Research Institute (<http://bacpac.chori.org/>) and grown in SW105 (ref. 32) cells. A single colony was expanded in LB broth with chloramphenicol at 32 °C. When the culture reached an attenuation (*D*) at 600 nm of 0.5, recombinering proteins were induced by incubation at 42 °C for 20 min. After induction, cells were spun at 5,000g, washed in ice-cold water, and resuspended in 200 μ l ice-cold 10% glycerol. The recombination cassette consisted of eGFP-FRT-PGKgb2-neo/kan-FRT, a 50-base-pair (bp) homology region in *LGR5*, and a 20-bp homology region to pEGFP-PGKneo. The homology regions were selected to replace the initiator methionine of *LGR5* with that of eGFP followed by a bovine growth hormone polyadenylation signal and an FRT-flanked bifunctional kanamycin/G418 resistance cassette. The recombination cassette was electroporated into SW105 cells, and cells were selected on plates with chloramphenicol and kanamycin (kan; 50 μ g ml⁻¹). Clones were analysed for properly targeted *LGR5* BAC by PCR (primer sequences listed in Methods) and confirmed by sequencing and nucleofected into single-cell suspensions of H9 ES cells using the Amaxa Human Stem Cell Nucleofector Starter Kit. Cells were selected for in G418 (200 ng ml⁻¹) for 2 weeks. G418-resistant cells were maintained in antibiotic indefinitely. All primer sequences are listed in the Methods.

Generation and characterization of iPS cell lines. Primary human foreskin fibroblasts (HFFs) were cultured from neonatal human foreskin tissue and obtained from two donors through the Department of Dermatology, University of Cincinnati, and were a gift from S. Wells. HFFs were cultured in fibroblast media consisting of DMEM (Invitrogen) supplemented with 10% FCS (Hyclone) and used for reprogramming between passages 5 and 8. EBNA1/OriP-based episomal plasmids pCLXE-hOct3/4-shp53, pCLXE-hSox2-Klf4, pCLXE-hLmyc-Lin28 and pCLXE-GFP used for this study were previously described³³ and obtained from Addgene (ID numbers: 27077, 27078, 27080 and 27082, respectively). The optimized Human Dermal

Fibroblast Nucleofector Kit (VPD-1001; Lonza) was used for transfection of HFFs with episomal plasmids. In brief, for each transfection 1 \times 10⁶ HFFs were pelleted by centrifugation at 200g for 10 min at room temperature, resuspended in 100 μ l Nucleofector Solution and nucleofected with 1.25 μ g each episomal plasmid (program U20). Cells from two transfections (2 \times 10⁶ total cells) were replated in a 10-cm tissue culture plate in fibroblast media, and cultured at 37 °C/5% CO₂. Six days after transfection, 4.5 \times 10⁵ HFFs were replated in fibroblast media in a gelatin-coated 10-cm dish containing 1.07 \times 10⁶ irradiated mouse embryonic fibroblasts. Starting on day 7 post-transfection, cells were fed daily with DMEM/F12 media supplemented with 20% knockout serum replacement, 1 mM L-glutamine, 0.1 mM β -mercaptoethanol, 0.1 mM non-essential amino acids, and 4 ng ml⁻¹ basic FGF (all from Invitrogen). Approximately two weeks later, discrete colonies with ES-cell-like morphology were manually excised and replated in mTesR1 media (Stem Cell Technologies) in tissue culture dishes coated with ES-cell-qualified Matrigel (Becton Dickinson). Following adaptation to mTesR1/Matrigel culture, iPS cells that maintained robust proliferation and ES-cell-like morphology with minimal spontaneous differentiation were expanded for characterization including testing for mycoplasma (MycopAlert kit, Lonza) and cryopreservation.

Standard metaphase spreads and G-banded karyotypes were determined by the CCHMC Cytogenetics Laboratory. For teratoma formation, iPS cells from three wells of a six-well dish were combined and gently resuspended in ice-cold DMEM/F12. Immediately before injection, Matrigel was added to a final concentration of ~33% and cells were injected subcutaneously into immunocompromised non-obese diabetic–severe combined immunodeficient (NOD scid gamma (NSG)) mice with IRB approval (3D06043). Tumours formed within 6–12 weeks. Excised teratomas were fixed, embedded in paraffin, and sections were stained with haematoxylin and eosin for histological examination. Taqman hPSC Scorecard Assay (Life Technologies) was performed according to manufacturer's instructions.

***Helicobacter pylori* infection.** *Helicobacter pylori* strain G27 (ref. 34) and a mutant G27 strain lacking CagA (Δ CagA)³⁵ were grown on blood agar plates consisting of Columbia Agar Base (Fisher Scientific), 5% horse blood (Colorado Serum Company), 5 μ g ml⁻¹ vancomycin and 10 μ g ml⁻¹ trimethoprim as described previously³⁶. For organoid injections, *H. pylori* were resuspended in brucella broth at a concentration of 1 \times 10⁹ bacteria per ml and loaded onto the Nanoject II (Drummond) microinjector apparatus. Approximately 200 nl (containing 2 \times 10⁵ bacteria) was injected directly in the lumen of each organoid, and injected organoids were cultured for 24 h. Brucella broth was injected as a negative control. Before all infection experiments, antibiotics were removed from the organoid growth medium.

Immunofluorescent staining. All tissues were fixed in 4% paraformaldehyde for either 1 h at room temperature for frozen processing or overnight at 4 °C for paraffin processing. Control mouse embryonic and postnatal tissues were obtained from time plugged CD-1 female mice with IRB approval (3D06043). For frozen sections, tissue was protected in 30% sucrose overnight at 4 °C, then embedded in OCT (Tissue-Tek), and cut at 10 μ m. For paraffin sections, tissue was processed through a graded ethanol series, followed by xylene, and then embedded in paraffin and cut at 7 μ m. Tissue culture cells were fixed for 15 min at room temperature and stained directly. For staining, frozen slides were thawed to room temperature and rehydrated in PBS, while paraffin slides were deparaffinized and subjected to antigen retrieval. Slides were blocked in 5% normal donkey serum (Jackson Immuno Research) in PBS plus 0.5% Triton-X for 30 min at room temperature. Primary antibodies (listed in the Methods) were diluted in blocking buffer and incubated overnight at 4 °C. Slides were washed in PBS and incubated with secondary antibody for 1 h at room temperature, and coverslips were mounted using Fluoromount-G (Southern Biotech). Confocal images were captured on a Nikon A1Rsi inverted confocal microscope.

Transmission electron microscopy imaging. Organoids were fixed in 2% glutaraldehyde plus 2% paraformaldehyde in 0.1 M sodium cacodylate buffer (pH 7.4) for 16 h at 4 °C. Organoids were then washed using 0.1 M sodium cacodylate buffer followed by a 1-h incubation using 4% osmium tetroxide, washed and then dehydrated using 25–100% ethanol (series of dilutions), embedded using propylene oxide/LX112. Blocks were sectioned (150 nm) and stained with 2% uranyl acetate followed by lead citrate. Tissue was visualized using a Hitachi transmission electron microscope equipped with an AMT Image Capture Engine version 5.42.366 and MicroFIRE by Optronics camera using AMTV600 digital camera software.

RNA isolation and qPCR. Total RNA was isolated from tissues using the Nucleospin RNA II kit (Machery-Nagel). Reverse transcription was performed from 100 ng RNA using Superscript VILO cDNA Synthesis Kit (Invitrogen) according to the manufacturer's protocol. qPCR was performed using Quantitect SybrGreen Master Mix (Qiagen) on a CFX-96 Real-time PCR Detection System (BioRad). Analysis was performed using the $\Delta\Delta C_t$ method. PCR primers were designed using sequences from qPrimerDepot (<http://primerdepot.nci.nih.gov>) and are listed in the Methods.

RNA sequencing. RNA library construction and RNA sequencing was performed by the University of Michigan DNA Sequencing Core, using an Illumina Hi-Seq 2000 platform. The UM Bioinformatics Core downloaded the read files from the

Sequencing Core, and concatenated those into a single .fastq file for each sample. Publicly available RNAseq data sets were downloaded from EBI-AE database (accession number E-MTAB-1733)³⁷ and NCBI-GEO (SRA) database (accession number GSE18927)³⁸. Raw reads were quality checked for each sample using FastQC (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/>) (version 0.10.1) to identify features of the data that may indicate quality problems (for example, low quality scores, over-represented sequences, inappropriate GC content). Initial quality control report indicated over-representation of Illumina adaptor sequences in samples from EBI-AE data set and NCBI-GEO data set. Adaptor sequences were trimmed from the reads using Cutadapt (version 0.9.5). We used the software package Tuxedo Suite for alignment, differential expression analysis, and post-analysis diagnostics^{39–41}. In brief, reads were aligned to the reference transcriptome (UCSC hg19, <http://genome.ucsc.edu>) using TopHat (version 2.0.9) and Bowtie (version 2.1.0.0). We used default parameter settings for alignment, with the exception of: ‘-b2-very-sensitive’ instructing the software to spend extra time searching for valid alignments, as well as ‘-no-coverage-search’ and ‘-no-novel-juncs’ to limit the read mapping to known transcripts. In addition, we used FastQC for a second round of quality control (post-alignment), to ensure that only high-quality data would be input to expression quantitation and differential expression analysis. We used Cufflinks/CuffDiff (version 2.1.1) for expression quantitation and differential expression analysis, using UCSC hg19.fa as the reference genome sequence and UCSC hg19.gtf as the reference transcriptome annotation. We generated diagnostic plots using the CummeRbund package (<http://compbio.mit.edu/cummeRbund/>).

Flow cytometry. Cells were incubated with Accutase solution at 37 °C until a single-cell suspension was obtained. Cells were washed with DMEM/F12 (Life Technologies) then spun down at 300g for 3 min. Cells were re-suspended in PBS and FACS of GFP-HI ES cells was performed on a FACSCalibur. At least 5×10^5 GFP-HI ES cells were collected for each sample ($n = 3$ biological replicates). After collection, cells were spun down, and immediately resuspended in lysis buffer for RNA isolation.

Immunoprecipitation and western blot analysis. *Helicobacter pylori*-infected organoids were collected from Matrigel in ice-cold PBS and centrifuged at 150g for 5 min. Tissue was lysed in M-PER Mammalian Protein Extract Reagent (Thermo Scientific) supplemented with protease inhibitors (Roche). Ten micrograms of total protein from the cell lysates was immunoprecipitated with anti-c-Met antibody (2 µg; Cell Signaling 4560) at 4 °C for 16 h. Protein A/G agarose beads (20 µl; Santa Cruz Biotechnology) were then added and the samples were incubated at 4 °C for 16 h. Immunoprecipitates were washed three times in PBS and then resuspended in Laemmli loading buffer containing β-mercaptoethanol (40 µl; BioRad). The positive control for phosphorylated c-Met was whole-cell lysate from EGF-stimulated A-431 epidermoid carcinoma cells (Santa Cruz Biotechnology), which co-migrates with the phosphorylated tyrosine band in c-Met-immunoprecipitated lysates. The positive control for immunoprecipitated CagA is lysate from G27 *H. pylori*, showing the CagA band co-migrates with the band in the c-Met-immunoprecipitated lysates. Samples were run on a 4–20% Tris-Glycine Gradient Gel (Invitrogen) and run at 80 V for 3.5 h. Gels were transferred to nitrocellulose membranes (Whatman Protran, 0.45 µm) at 105 V for 1.5 h. Membranes were blocked in KPL Detector Block Solution (Kirkegaard & Perry Laboratories) for one hour at room temperature and then incubated with primary antibody overnight at 4 °C. Primary antibodies used: anti-phosphotyrosine (Santa Cruz, sc-7020; 1:100); anti-c-Met (Abcam, ab59884; 1:100); and anti-*H. pylori* CagA (Abcam, ab90490; 1:100). Membranes were washed and incubated in Alexa Fluor anti-mouse 680 (Invitrogen; 1:1000) secondary antibody. Blots were imaged using the Odyssey Infrared Imaging Software System (Licor).

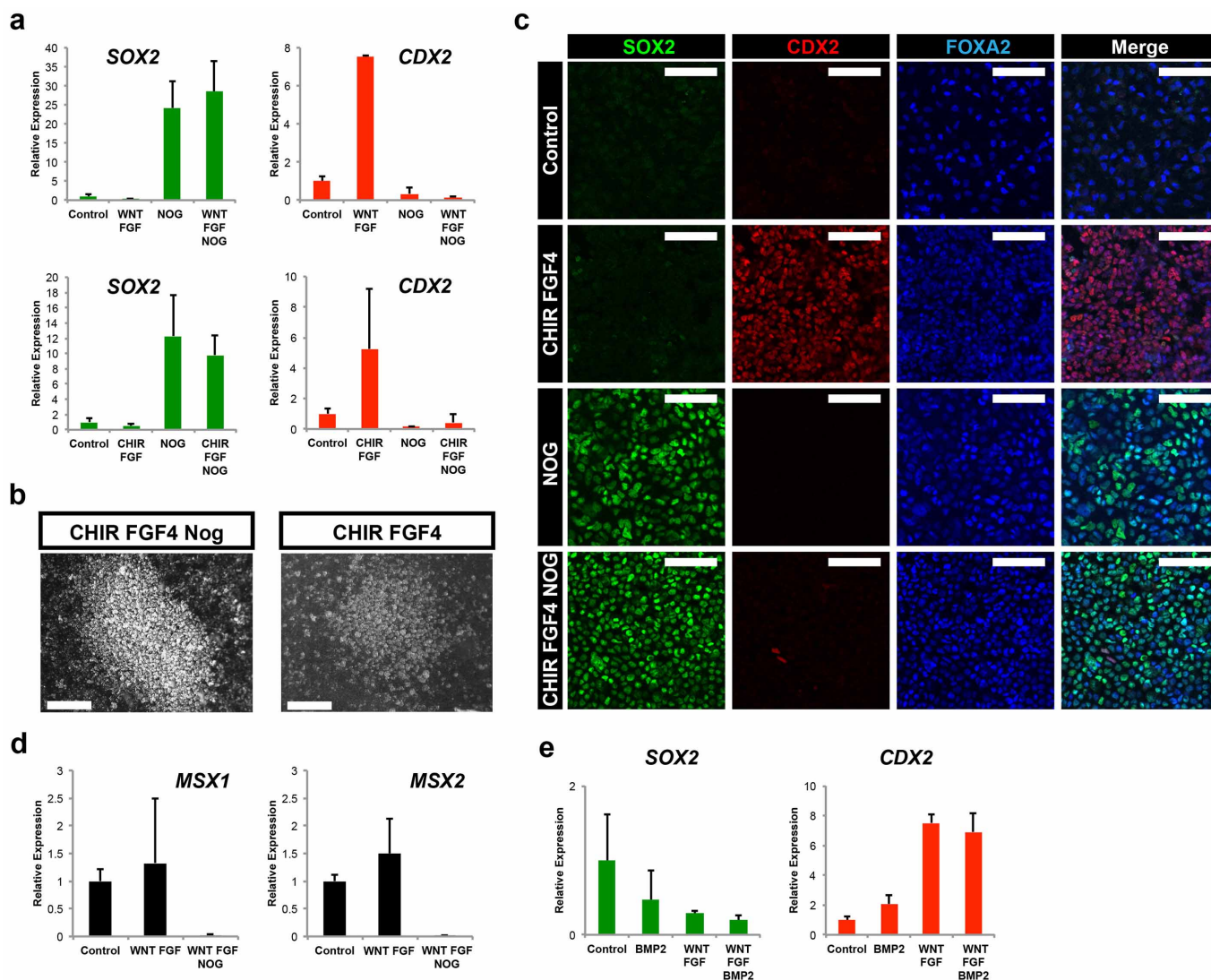
Primary antibodies used for immunofluorescent staining. The primary antibodies used in immunofluorescence staining are listed below with target, species, company, catalogue number and dilution. Acta2, rabbit, GeneTex, GTX100034, 1:200; Atp4b, mouse, Thermo Fisher, MA3-923, 1:1,000; aPKC, rabbit, Santa Cruz, sc216, 1:200; β-Catenin, rabbit, Santa Cruz, sc7199, 1:100; Cdx2, mouse, Biogenex, MU392A-UC, 1:500; Chga, rabbit, Immunostar, 20086, 1:500; Desmin, goat, Santa Cruz, sc7559, 1:200; E-cadherin, mouse, BD Biosciences, 610182, 1:500; E-cadherin, goat, R&D Systems, AF648, 1:500; FoxF1, goat, R&D Systems, AF4798, 1:500; Gastrin, rabbit, Dako, A0568, 1:1,000; Gata4, mouse, Santa Cruz, sc25310, 1:200; GFP, rabbit, Invitrogen, A11122, 1:1,000; Ghrelin, goat, Santa Cruz, sc10368, 1:200; *H. pylori*, rabbit, Abcam, ab80519, 1:1,000; Hnf1β, mouse, BD Biosciences, 612504, 1:500; Hnf1β, goat, Santa Cruz, sc4711, 1:500; Ki67, rabbit, Abcam, ab833, 1:200; Ki67, rat, Dako, m7249, 1:100; Klf5, rat, gift from R. Nagai and T. Shindo, 1:2,000; Muc5AC, mouse, Abcam, ab3649, 1:500; Muc6, mouse, Abcam, AB49462, 1:100; Nanog, rabbit, Abcam, ab21624, 1:500; Oct3/4, mouse, Santa Cruz, sc5279, 1:500; Pdx1, goat, Abcam, ab47383, 1:5,000; pHH3, rabbit, Cell Signaling, 9701, 1:500; Serotonin (5-HT), rabbit, Immunostar, 20080, 1:1,000; Somatostatin, goat, Santa Cruz, sc7819, 1:100; Sox2, goat, Santa Cruz, sc17320, 1:500; Sox2, rabbit, Seven Hills Bioreagents, WRAB-1236, 1:1,000; Sox9, rabbit, Millipore, AB5535, 1:10,000; Tff2, goat, Santa Cruz, sc23558, 1:500; Vimentin, goat, Santa Cruz, sc7557, 1:200.

Primer sequences. The primers used for qPCR analyses were: *ATP4A*, forward 5'-TGTTAGTAGCCAAAGCAGCC-3', reverse 5'-TGCCATCCAGGCTAGTGAG-3'; *ATP4B*, forward 5'-ACCACGTAGAAGGCCACGT-3', reverse 5'-TGGAGGAGTTCACAGCGTTAC-3'; *BAPX1*, forward 5'-CAACACCGTCGTCCTCG-3', reverse 5'-CCGCTTCCAAAGACCTAGAG-3'; *CDX2*, forward 5'-CTGGAGCTGGAGAAGGAGTTTC-3', reverse 5'-ATTTTAACCTGCCTCTCAGAGAGC-3'; *CHGA*, forward 5'-TGACCTCAACGATGCATTTTC-3', reverse 5'-CTGTCTGGCTCTCTGCTC-3'; *GAPDH*, forward 5'-CCCATCACCATTCTTCCA GGAG-3', reverse 5'-CTTCTCCATGGTGGTGAAGACG-3'; *GAST*, forward 5'-CAGAGCCAGTGC AAAGATCA-3', reverse 5'-AGAGACCTGAGAGGCACCAG-3'; *GHRL*, forward 5'-GCTGGTACTGAACCCCTGAC-3', reverse 5'-GATGGAGTCAAGCAGAAGG-3'; *GKN1*, forward 5'-AGCTAGGGCAGGAGCTAGAAA-3', reverse 5'-GCTTGCTACTCTCTGTGCC-3'; *HNF1B*, forward 5'-TCACAGATACCAGCAGCATCAGT-3', reverse 5'-GGGCATCACCAGGCTTGT-3'; *HNF6*, forward 5'-TGTTGCTCTATCCTTCCCA-3', reverse 5'-GGA GGATGTGGAAGTGCT-3'; *MIST1*, forward 5'-TGCTGGACATGGTCAGGAT-3', reverse 5'-CGGACAAGAAGCTCTCCAAG-3'; *MSX1*, forward 5'-GGTTCGTCTGTGTTTGGC-3', reverse 5'-CCCGAGAAGCCCGAGAG-3'; *MSX2*, forward 5'-GGTCTGTGTTTCTCCTCAGGG-3', reverse 5'-AAATTCAGAAGATGGAGCGG-3'; *MUC2*, forward 5'-TGTAGGCATCGCTCTTCTCA-3', reverse 5'-GACACCATCTACCTCACCCG-3'; *MUC5AC*, forward 5'-CCAAGGAGAACTCCCATAT-3', reverse 5'-CCAAGCTATCTCTGAG-3'; *MUC6*, forward 5'-CAGCAGAGGAGATACGTTCAAG-3', reverse 5'-GTGGGTGTTTCTTGCTGTGTCATC-3'; *NEUROG3*, forward 5'-CTTCGTCTTCCGAGGCTCT-3', reverse 5'-CTATCTTTTGGCCCGGTAG-3'; *PDX1*, forward 5'-CGTCCGCTTGTCTCTCTC-3', reverse 5'-CCTTTCCTATGGATGAAGTC-3'; *PTF1A*, forward 5'-AGAGAGTGTCTGCTAGGGG-3', reverse 5'-CCAGAAGGTCATCATCTGCC-3'; *SST*, forward 5'-GCGCTGTCCATCGTCTCTGGCCC-3', reverse 5'-AGCCGGTTTGTAGTTAGCAGAT-3'; *SOX2*, forward 5'-GCTTAGCCTCGTCGATGAAC-3', reverse 5'-AACCCCAAGATGCACAACCTC-3'; *TFF1*, forward 5'-AATTCTGTCTTTCACGGGGG-3', reverse 5'-GGAGAACAAGGTGATCTGC-3'; *TFF2*, forward 5'-TCTGAGACCTCCATGACGC-3', reverse 5'-ATGGATGCTGTTTCGACTCC-3'; *TFF3*, forward 5'-CACTCCTTGGGGGTGACA-3', reverse 5'-CTCCAGCTCTGCTGAGGAGT-3'.

The primers for *LGR5-eGFP* BAC cloning were: recombination cassette primers, forward 5'-GGTGCTGCTCTCCGCCCGCTCCGGCTCGTGGCCCCCTACTTCGGGCACCATGGTGAGCAAGGGCAGGA-3', reverse 5'-TTCCTTCCCTCTTAGTCTCTCTCCCGGAGTGACGTGGGGGAAGTACTTACCTATACGAAGTTATAAGCTT-3'.

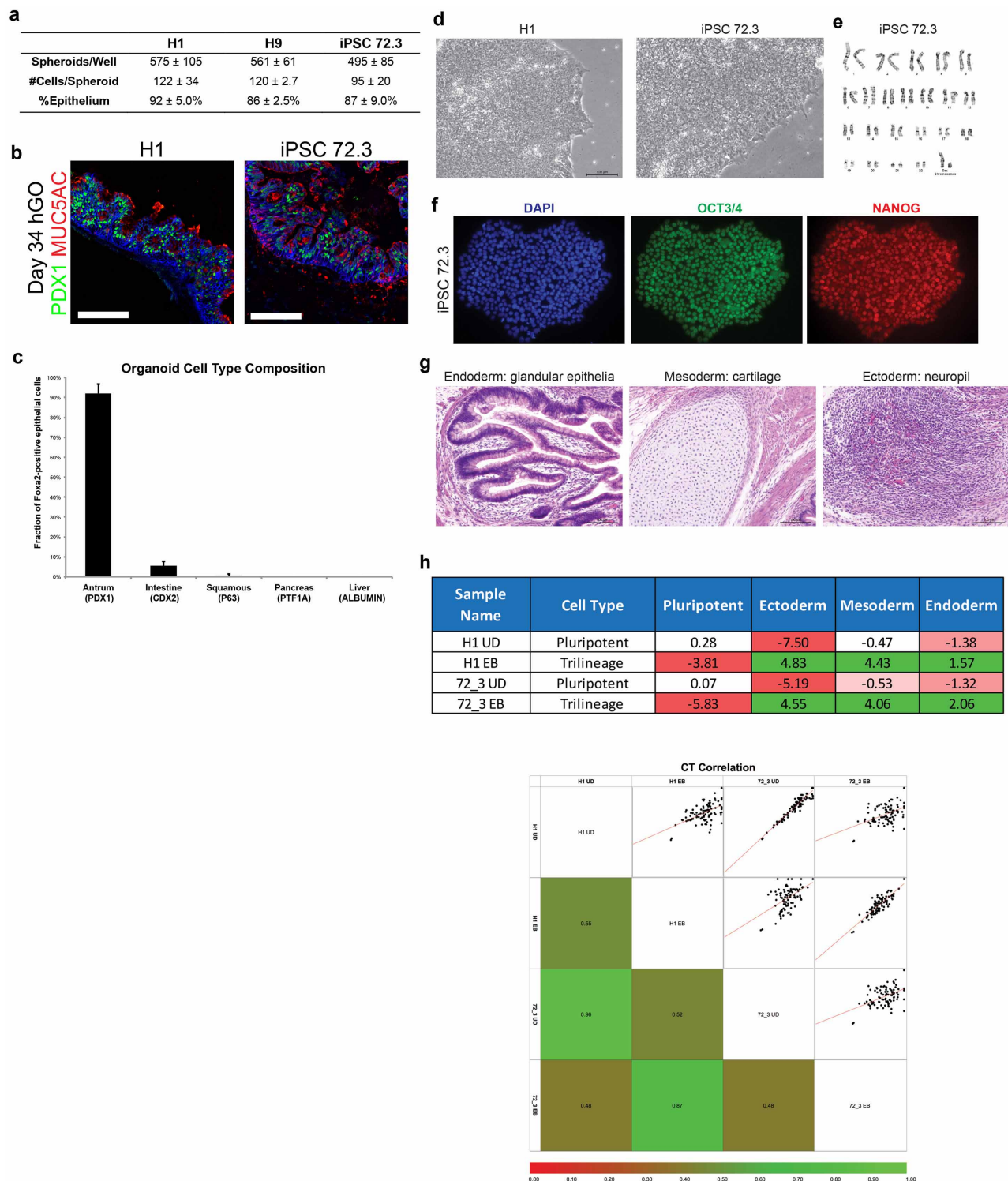
Colony screening no. 1, forward 5'-AGACGCCCGCTGAGTTGCGAG-3', reverse 5'-TGCACGCCGTAGGTCAGGGT-3'; colony screening no. 2, forward 5'-CAGCAGCGCTTCCCGGGT-3', reverse 5'-GGCGAATGGGCTGACCGCTT-3'.

30. Teo, A. K. K. *et al.* Activin and BMP4 synergistically promote formation of definitive endoderm in human embryonic stem cells. *Stem Cells* **30**, 631–642 (2012).
31. Meerbrey, K. L. *et al.* The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proc. Natl Acad. Sci. USA* **108**, 3665–3670 (2011).
32. Warming, S., Costantino, N., Court, D. L., Jenkins, N. A. & Copeland, N. G. Simple and highly efficient BAC recombineering using *galK* selection. *Nucleic Acids Res.* **33**, e36 (2005).
33. Okita, K. *et al.* An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells. *Stem Cells* **31**, 458–466 (2013).
34. Covacci, A. *et al.* Molecular characterization of the 128-kDa immunodominant antigen of *Helicobacter pylori* associated with cytotoxicity and duodenal ulcer. *Proc. Natl Acad. Sci. USA* **90**, 5791–5795 (1993).
35. Amieva, M. R., Salama, N. R., Tompkins, L. S. & Falkow, S. *Helicobacter pylori* enter and survive within multivesicular vacuoles of epithelial cells. *Cell. Microbiol.* **4**, 677–690 (2002).
36. Schumacher, M. A. *et al.* Gastric Sonic Hedgehog acts as a macrophage chemoattractant during the immune response to *Helicobacter pylori*. *Gastroenterology* **142**, 1150–1159 (2012).
37. Fagerberg, L. *et al.* Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteomics* **13**, 397–406 (2014).
38. Bernstein, B. E. *et al.* The NIH roadmap epigenomics mapping consortium. *Nature Biotechnol.* **28**, 1045–1048 (2010).
39. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
40. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
41. Trapnell, C. *et al.* Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnol.* **31**, 46–53 (2013).
42. Xie, R. *et al.* Dynamic chromatin remodeling mediated by polycomb proteins orchestrates pancreatic differentiation of human embryonic stem cells. *Stem Cells* **12**, 224–237 (2013).



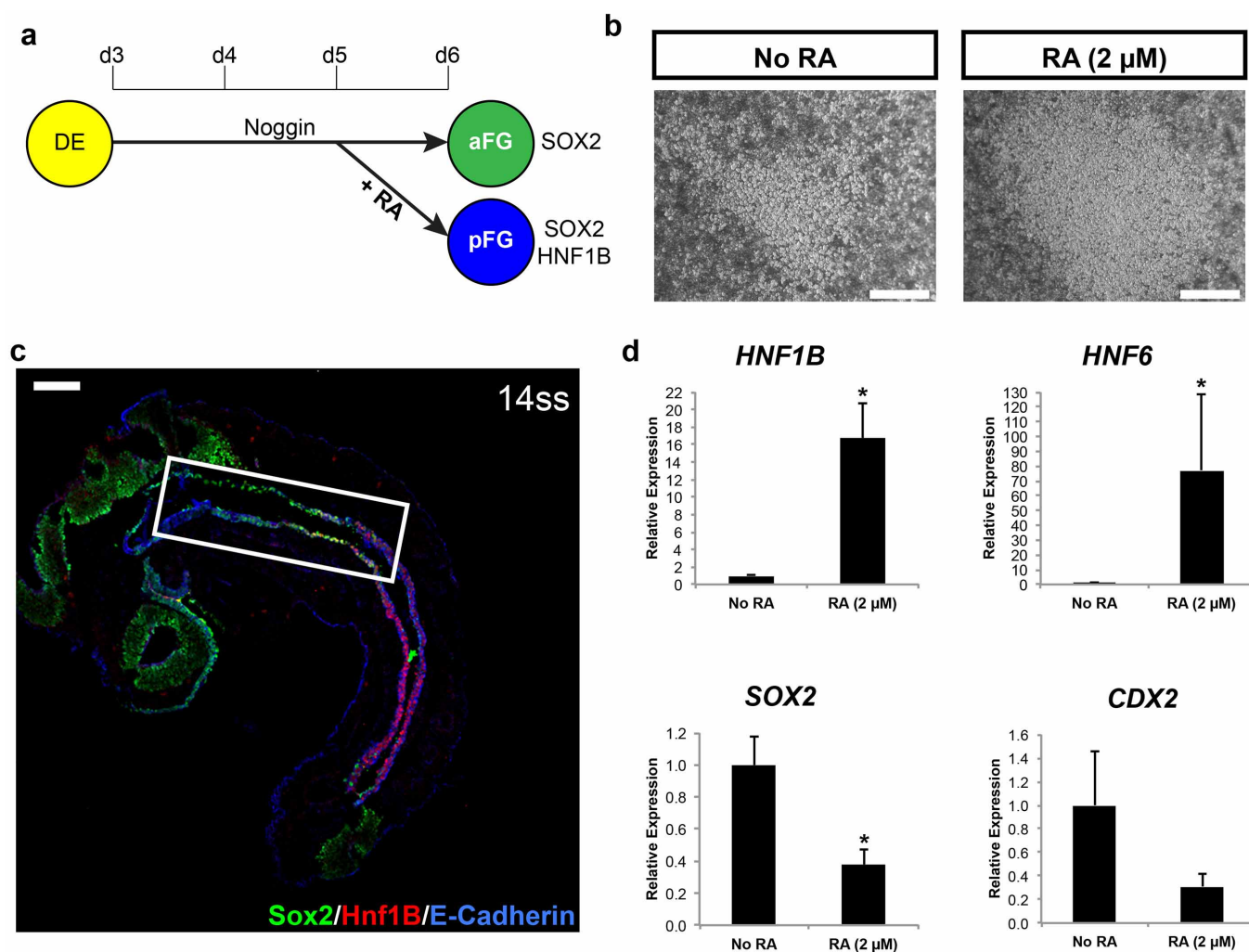
Extended Data Figure 1 | BMP signalling is required in parallel with activation of WNT and FGF to promote a posterior fate. **a**, Activation of WNT signalling with WNT3A or the GSK3 β inhibitor CHIR99021 (CHIR; 2 μ M) induced a posterior fate and this was blocked by BMP inhibition. $n = 3$ biological replicates per condition. **b**, Activation of WNT signalling with WNT3A (not shown) or CHIR induced gut tube morphogenesis and spheroid production. **c**, Immunofluorescent staining of monolayer cultures confirmed the high efficiency of CDX2 induction by CHIR/FGF treatment, and that NOG blocked posterior CDX2 expression and induced expression of the foregut

marker SOX2. **d**, qPCR analysis of BMP target genes *MSX1/2* indicated that BMP activity is not increased in response to WNT/FGF, but target genes are suppressed in response to NOG, suggesting that NOG acts on endogenous BMP signalling. $n = 3$ biological replicates per condition. **e**, Addition of BMP2 (100 ng ml $^{-1}$) did not substitute for or augment the ability of WNT/FGF to posteriorize endoderm. These data indicate that the posteriorizing effect of WNT/FGF is not mediated by upregulation of BMP signalling but does require endogenous BMP activity. $n = 3$ biological replicates per condition. Scale bars, 1 mm (**b**) and 100 μ m (**c**). Error bars represent s.d.



Extended Data Figure 2 | Gastric organoid differentiation is efficient in multiple pluripotent stem cell lines. **a**, Table comparing spheroid formation and characteristics between two human ES cell lines (H1 and H9) and one iPS cell line (72.3). Spheroid number was averaged from $n = 8$ wells per cell line; total cells per spheroid and epithelial composition were determined from whole mount staining (DAPI for total cell number and FOXA2 for epithelial cells) and quantification from $n = 6$ spheroids per cell line. Error bars represent s.d. **b**, Immunofluorescent staining of day-34 hGOs derived from ES cell line H1 and iPS cell line 72.3. iPS-cell-derived organoids exhibit the same morphological and molecular features of those derived from ES cells. **c**, Organ epithelial cell type quantification in day-34 hGOs. Greater than 90% of the epithelium is antral, indicated by PDX1 expression and lack of PTF1A expression, whereas less than 5% express markers associated with other organs

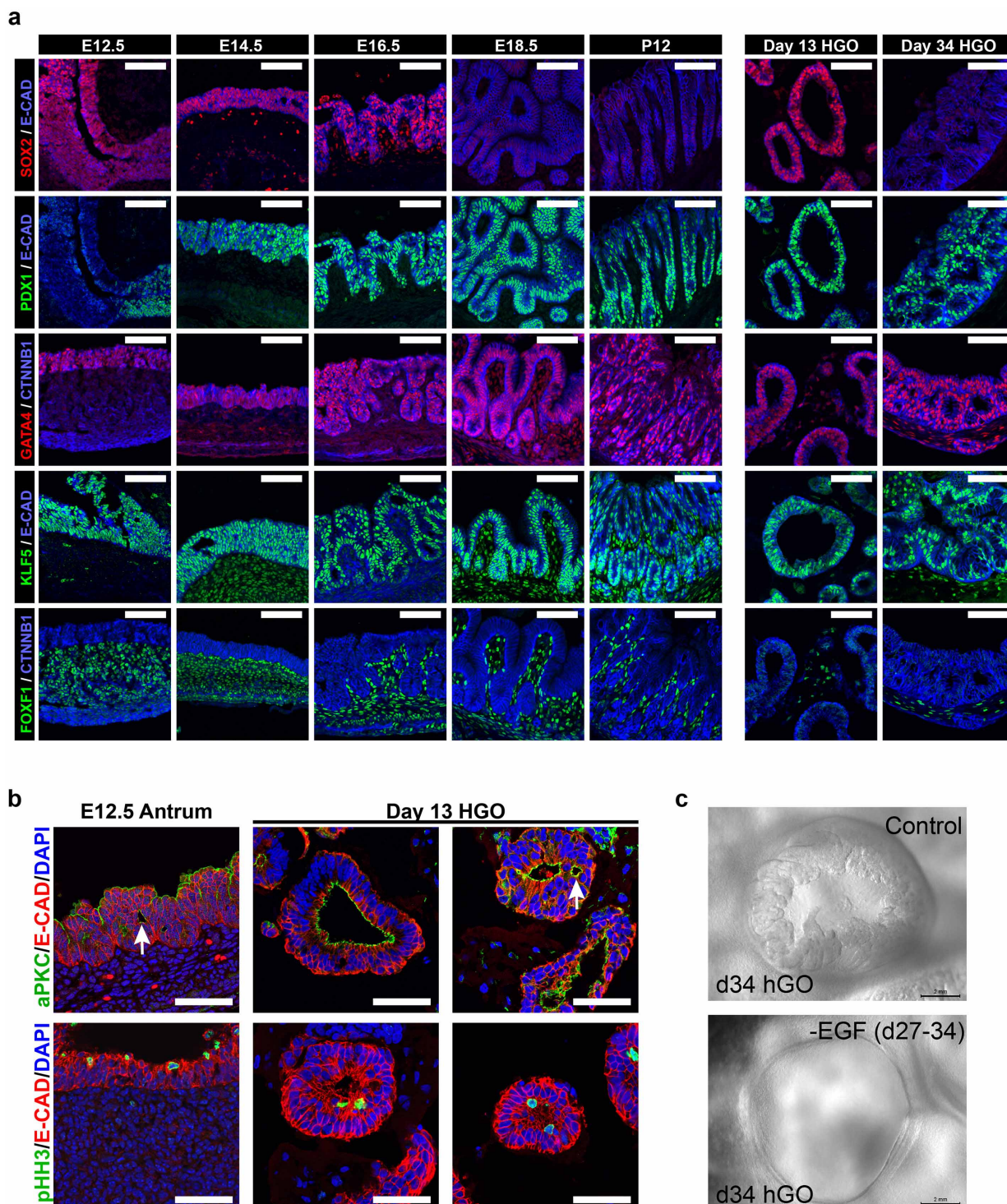
derived from endoderm including CDX2 (intestine), albumin (liver) and p63 (squamous epithelium). Data shown are averages from $n = 6$ hGOs. **d–g**, Characterization of iPS cell line 72.3 used in **a**. **d**, e, iPS cell line 72.3 exhibited normal morphological characteristics of pluripotent stem-cell colonies, as compared to the H1 hESC line (**d**) and had a normal 46;XY karyotype (**e**). **f**, g, iPS cell line 72.3 expressed pluripotent markers OCT3/4 and NANOG (**f**), and demonstrated pluripotency by differentiation into endoderm, mesoderm, and ectoderm lineages in an *in vivo* teratoma assay (**g**). **h**, Human pluripotent stem-cell scorecard assay results demonstrating that ES cell line H1 and iPS cell line 72.3 have similar pluripotency and differentiation potential, and that iPS cell line 72.3 does not have a lineage bias. EB, differentiated as embryoid bodies for 14 days; UD, undifferentiated. Scale bars, 100 μ m. Error bars represent s.d.



Extended Data Figure 3 | Retinoic acid posteriorizes foregut endoderm.

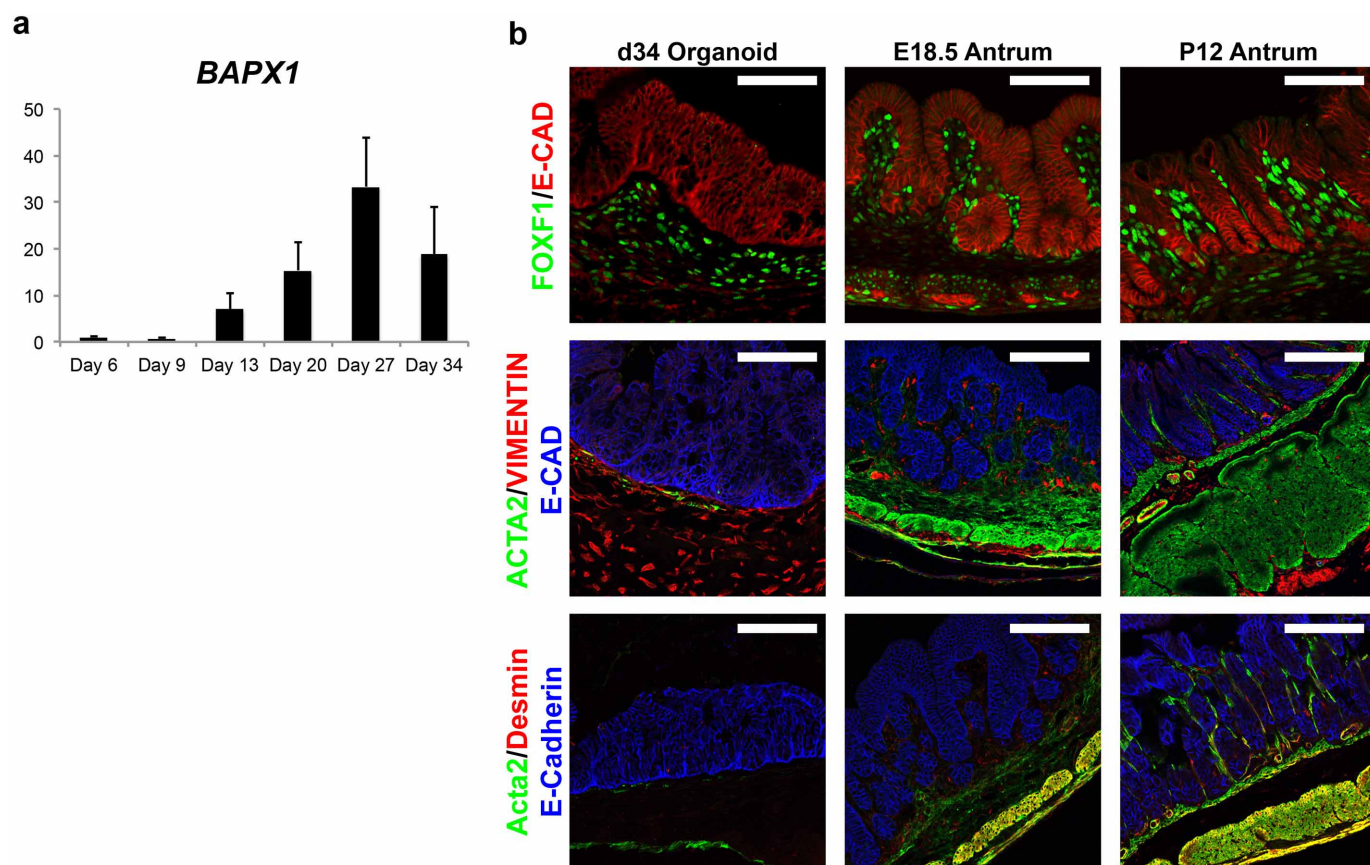
a, Lineage diagram that summarizes the patterning effects of noggin and retinoic acid in the formation of both anterior and posterior foregut endoderm (aFG and pFG, respectively). **b**, Bright-field images show that retinoic acid increased the number of spheroids that are produced from foregut monolayer cultures. **c**, A lower power image of Fig. 1d showing immunofluorescent image of an E8.5, 14-somite stage embryo with Hnf1 β protein localized to the posterior portion of the foregut. Boxed region of embryo is shown in Fig. 1d.

d, qPCR analysis of gene expression in foregut spheroids treated with retinoic acid. Posterior foregut markers *HNF1B* and *HNF6* were robustly induced by 24-h exposure to retinoic acid. Although retinoic acid induced posterior foregut gene expression it did not induce expression of the posterior marker *CDX2*. * $P < 0.05$; Student's *t*-test; $n = 3$ biological replicates per condition, data representative of 3 independent experiments. Scale bars, 1 mm (**b**) and 100 μ m (**c**). Error bars represent s.d.



Extended Data Figure 4 | hGOs recapitulate normal antrum development of mouse embryos. **a**, Comparison of transcription factor expression between hGO development and *in vivo* stomach organogenesis. Four embryonic stages (E12.5, E14.5, E16.5 and E18.5) and one postnatal stage (P12) of *in vivo* antrum development were analysed for expression of the following transcription factors: Sox2, Pdx1, Gata4, Klf5 and FoxF1. The same markers were analysed at two stages (days 13 and 34) of *in vitro* hGO development and revealed that organoid development parallels that which occurs *in vivo*. At early stages of antrum development the epithelial marker Sox2 was expressed ubiquitously but at later stages it is downregulated, while other epithelial transcription factors, Pdx1, Gata4 and Klf5, exhibit persistent expression in the epithelium throughout development. Both early- and late-stage hGOs contain FoxF1⁺ mesenchymal cells surrounding the epithelium, similar to the *in vivo* antrum. **b**, Early-stage hGOs exhibit stereotypic epithelial architecture and nuclear

behaviour. At day 13, hGOs contained pseudo-stratified epithelia that display apicobasal polarity marked by the apical marker aPKC and the basolateral marker E-cadherin, similar to the E12.5 mouse antrum. Furthermore, extensions of apical membrane (white arrows) were seen within deeper portions of the organoid epithelium. Both the E12.5 mouse antrum and day-7 hGOs appeared to undergo interkinetic nuclear migration, indicated by the presence of mitotic nuclei, phosphohistone H3 (pHH3), in only the apical portions of cells. **c**, EGF is required for morphogenesis in gastric organoids. Bright-field images demonstrate the requirement for EGF in epithelial morphogenesis including folding and gland formation at late stages of hGO differentiation. When EGF is removed from the growth medium at day 27, before glandular morphogenesis, the hGO epithelium retains a simple, cuboidal structure that fails to form glands. Scale bars, 100 μ m (**a**), 50 μ m (**b**) and 2 mm (**c**).

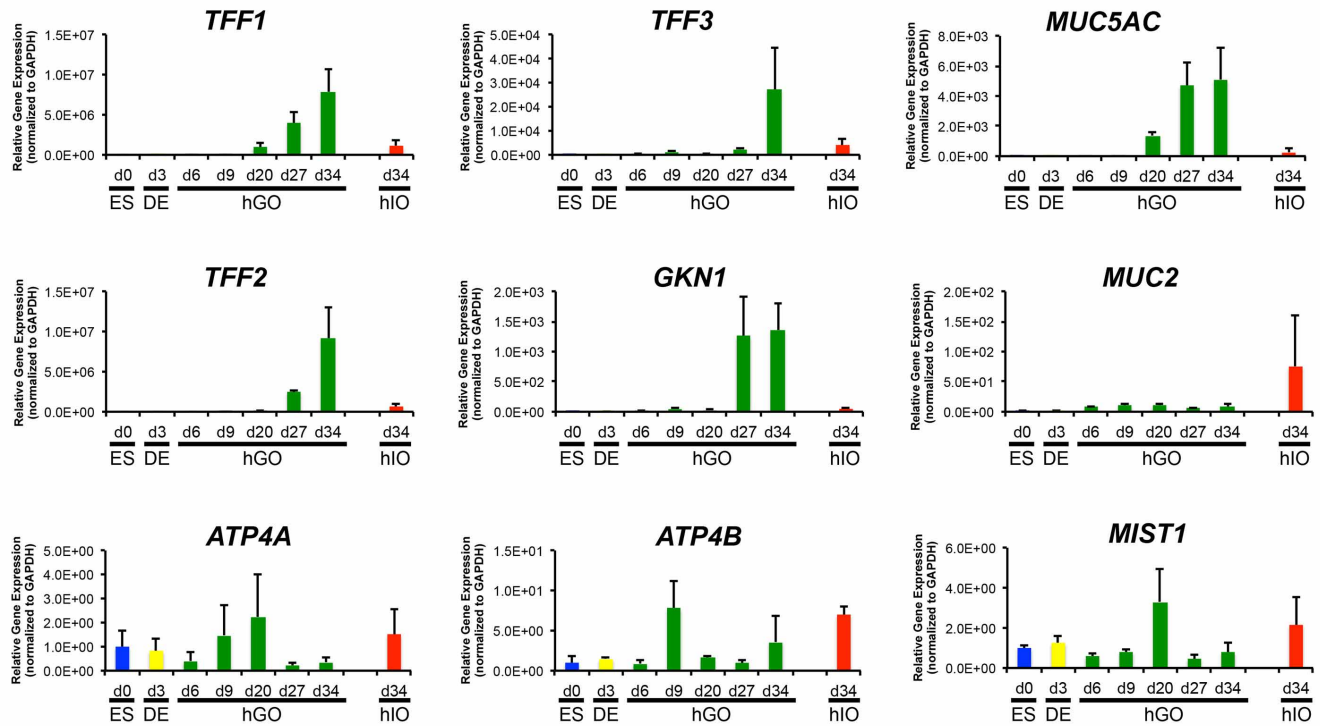


Extended Data Figure 5 | Mesenchymal differentiation in gastric organoids.

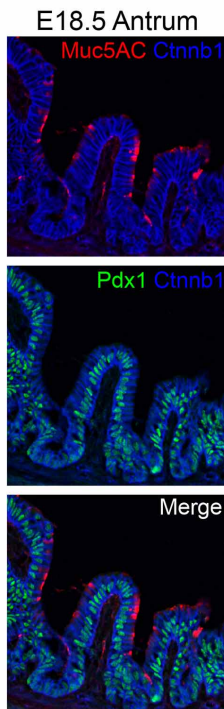
a, Temporal expression analysis of the antral mesenchyme transcription factor BAPX1. Similar to its known embryonic expression pattern, BAPX1 is upregulated during the earlier stages of hGO differentiation and then downregulated coincident with functional cell type marker expression. $n = 3$ biological replicates per time point. **b**, Staining for mesenchymal cell type

markers revealed that day-34 hGOs contain FOXF1/VIM-positive submucosal fibroblasts and a small number of VIM/ACTA2-expressing subepithelial fibroblasts. hGOs lack a robust smooth muscle layer, indicated by ACTA2/desmin-positive cells in the *in vivo* antrum. Scale bars, 100 μ m. Error bars represent s.d.

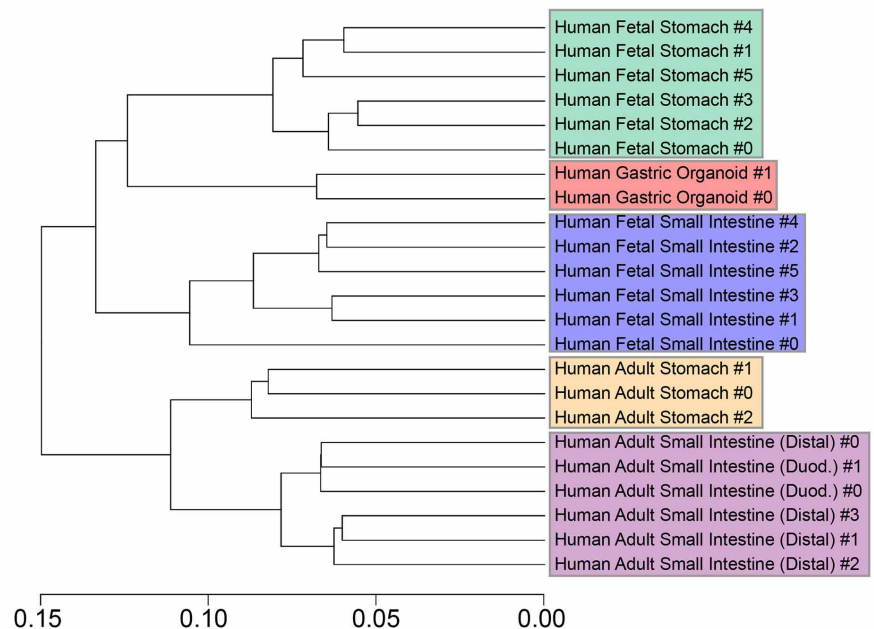
a



b

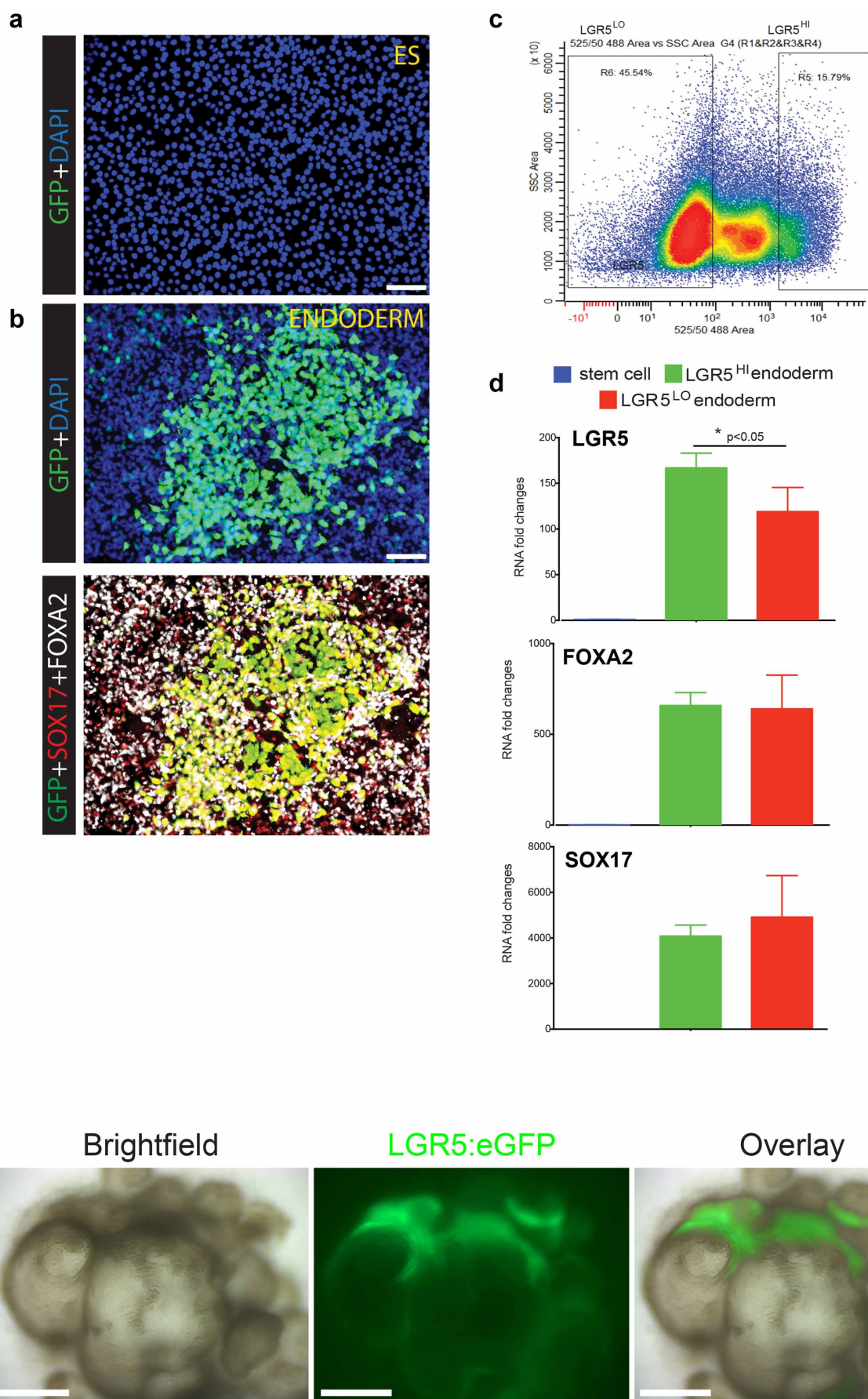


c



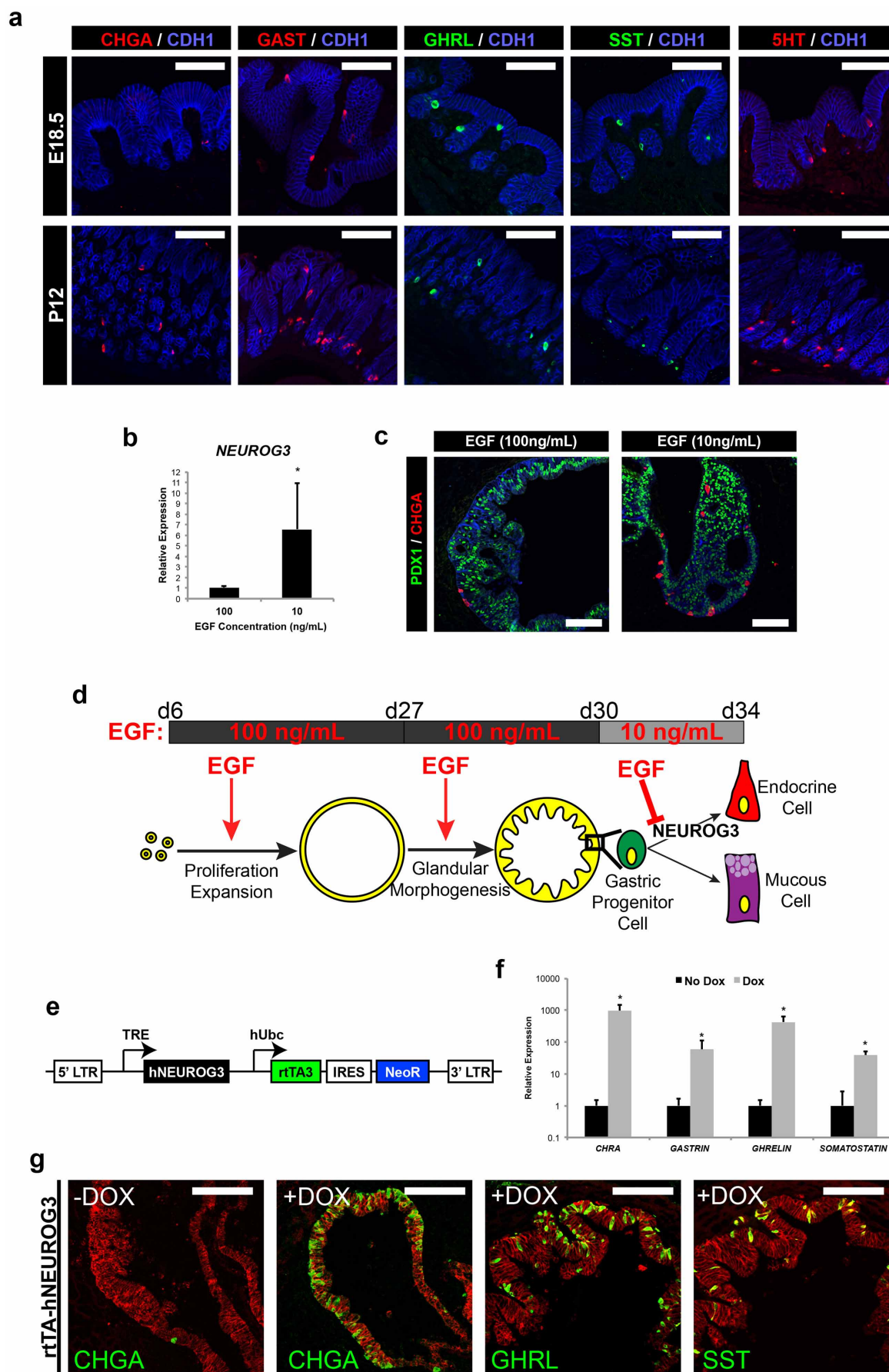
Extended Data Figure 6 | Induction of genes during development of hGOs that mark specific differentiated antral cell types. **a**, qPCR analyses of cell lineage differentiation marker expression at several stages throughout the gastric organoid differentiation protocol (days 0, 3, 6, 9, 20, 27 and 34) and day-34 human intestinal organoids (hIO). Beginning at day 27, hGOs robustly induced genes expressed in differentiated cell types including surface mucous cells (*MUC5AC*, *TFF1*, *TFF3* and *GKN1*) and antral gland cells (*TFF2*). hGOs do not upregulate the expression of markers found in fundic lineages such as parietal cells (*ATP4A* and *ATP4B*) and chief cells (*MIST1*) or intestinal goblet cells (*MUC2*). Expression levels are normalized to day-3 definitive endoderm

cultures. $n = 3$ biological replicates per time point. **b**, *Muc5AC*-expressing surface mucous cells in the late fetal (E18.5) mouse antrum are not yet confined to a pit region and are more broadly distributed through the antral epithelium. Furthermore, these pit cells exhibit high amounts of intracellular mucin staining, similar to day-34 hGOs. **c**, Global gene expression profiling of day-34 hGOs was performed using RNA-seq, and data were compared to published RNA-seq data sets from human tissues. Hierarchical clustering revealed that hGOs closely resemble human fetal stomach tissue but not human fetal intestine. Error bars represent s.d.



Extended Data Figure 7 | Characterization of *LGR5-eGFP* BAC transgenic reporter ES cell line. **a**, H9 *LGR5-eGFP* ES cell line did not show eGFP fluorescence in undifferentiated, pluripotent stem cells. **b**, Upon differentiation to definitive endoderm, robust eGFP expression was observed, consistent with published microarray and RNA-sequencing analyses that show *LGR5* as a highly enriched endoderm transcript^{6,42}. Top, DAPI and eGFP staining; bottom, eGFP co-localization with endoderm markers *SOX17* and *FOXA2*. **c**, FACS was used to sort *LGR5-eGFP^{LO}* and *LGR5-eGFP^{HI}* from 3-day activin-A-treated definitive endoderm cultures. **d**, qPCR was used to measure *LGR5*, *FOXA2* and *SOX17* expression levels in undifferentiated H9 *LGR5-eGFP* cells (blue bars, stem cell) and in FACS-purified H9 *LGR5-eGFP* endoderm (red bars, *LGR5-eGFP^{LO}*; green bars, *LGR5-eGFP^{HI}*). As expected, *LGR5*, *FOXA2*

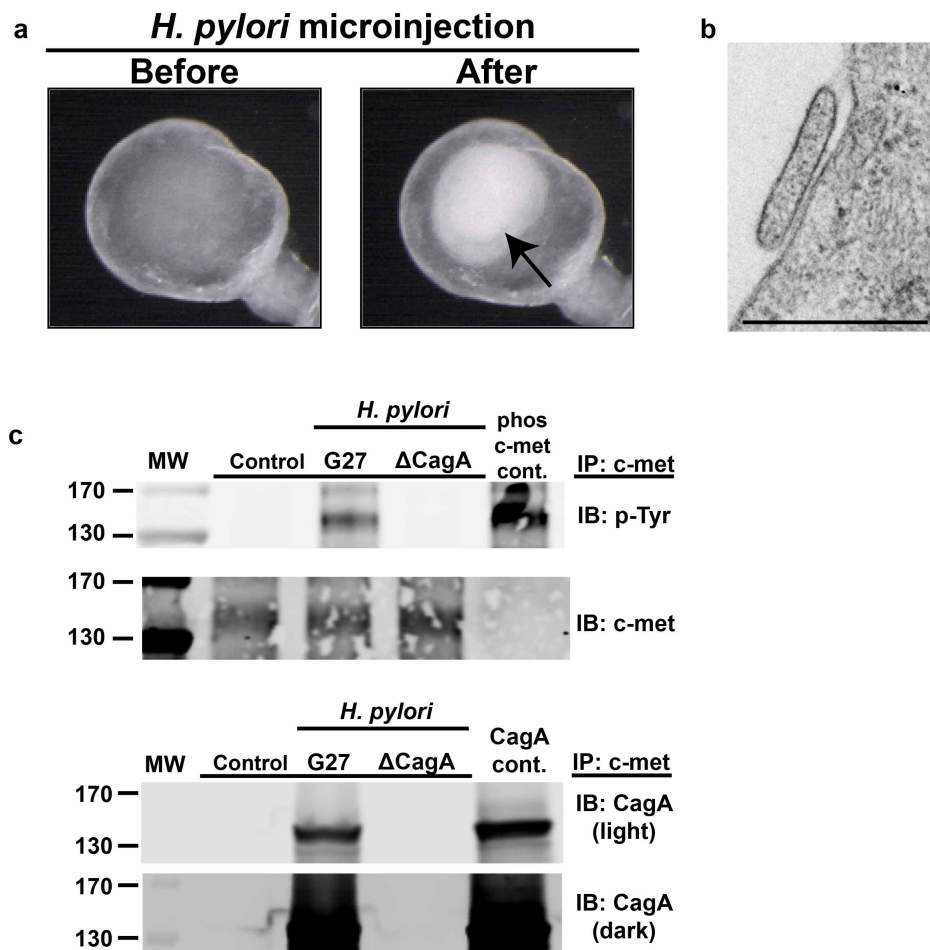
and *SOX17* were all highly enriched in both *LGR5-eGFP^{LO}* and *LGR5-eGFP^{HI}* endoderm populations compared to undifferentiated controls, and the *LGR5-eGFP^{HI}* cells showed significant enrichment of *LGR5* mRNA, but not *FOXA2* or *SOX17*, compared to the *LGR5-eGFP^{LO}* population. $n = 3$ biological replicates for each group and error bars represent s.e.m. $*P < 0.05$ using two-tailed Student's *t*-test. This analysis suggests that the *LGR5-eGFP* BAC construct drives eGFP expression in endoderm cells with the highest levels of *LGR5* expression. **e**, H9 *LGR5-eGFP* ES cells were differentiated into antral gastric organoids. Bright-field and eGFP stereomicrographs of day-30 hGOs showed that the organoid epithelium developed regionally-restricted areas of *LGR5-eGFP* expression, suggesting that *LGR5⁺* stem-cell populations formed during the differentiation of the organoids. Scale bars, 100 μ m.



Extended Data Figure 8 | *NEUROG3* expression and endocrine differentiation are reduced in a high EGF environment.

a, Endocrine cell differentiation in the antrum is first evident at E18.5 and highly robust at postnatal stages (P12 shown). As early as E18.5, all expected gastric endocrine subtype hormones are present, including gastrin, ghrelin, somatostatin and serotonin (5-HT). **b**, High levels of EGF (100 ng ml^{-1}) repressed *NEUROG3* expression, however a reduction in EGF concentration (10 ng ml^{-1}) at day 30 resulted in a significant increase in *NEUROG3* expression measured at day 34 by qPCR. $*P < 0.05$; Student's *t*-test; $n = 5$ biological replicates, data representative of 3 independent experiments. **c**, hGOs maintained in high concentrations of EGF (100 ng ml^{-1}) had very few endocrine cells at day 34, shown by staining for the pan-endocrine marker CHGA. However, a reduction of EGF concentration (to 10 ng ml^{-1}) at day 30 resulted in more physiological numbers of endocrine cells in the gastric epithelium. **d**, Schematic indicating the effects of EGF at different stages of hGO growth, morphogenesis, and cell

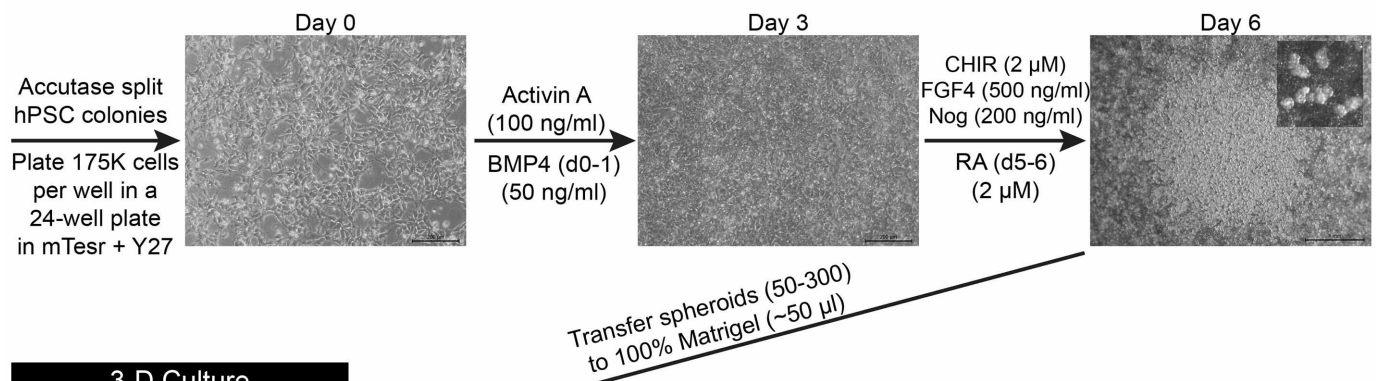
type specification. High levels of EGF were required at early developmental stages for growth and morphogenesis, however, it repressed endocrine differentiation at late stages of development; thus, the EGF concentration was reduced at day 30 to allow for endocrine cell development. **e**, To test whether EGF repression of endocrine differentiation occurs upstream of *NEUROG3*, hGOs were generated from an ES cell line stably transfected with a dox-inducible *NEUROG3*-overexpressing transgene. hGOs were maintained in high EGF (100 ng ml^{-1}), then at day 30 were treated with doxycycline ($1 \mu\text{g ml}^{-1}$) for 24 h and then analysed at day 34. **f**, **g**, Dox-treated hGOs show robust activation of endocrine markers *CHGA*, *GAST*, *GHRL* and *SST* (**f**), and they contain CHGA-, GHRL- and SST- expressing cells with endocrine morphology (**g**). $*P < 0.05$; Student's *t*-test; $n = 3$ biological replicates per condition, data representative of 2 independent experiments. Therefore, *NEUROG3* overexpression was sufficient to induce gastric endocrine cell fate in a high-EGF environment. Scale bars, $100 \mu\text{m}$. Error bars represent s.d.



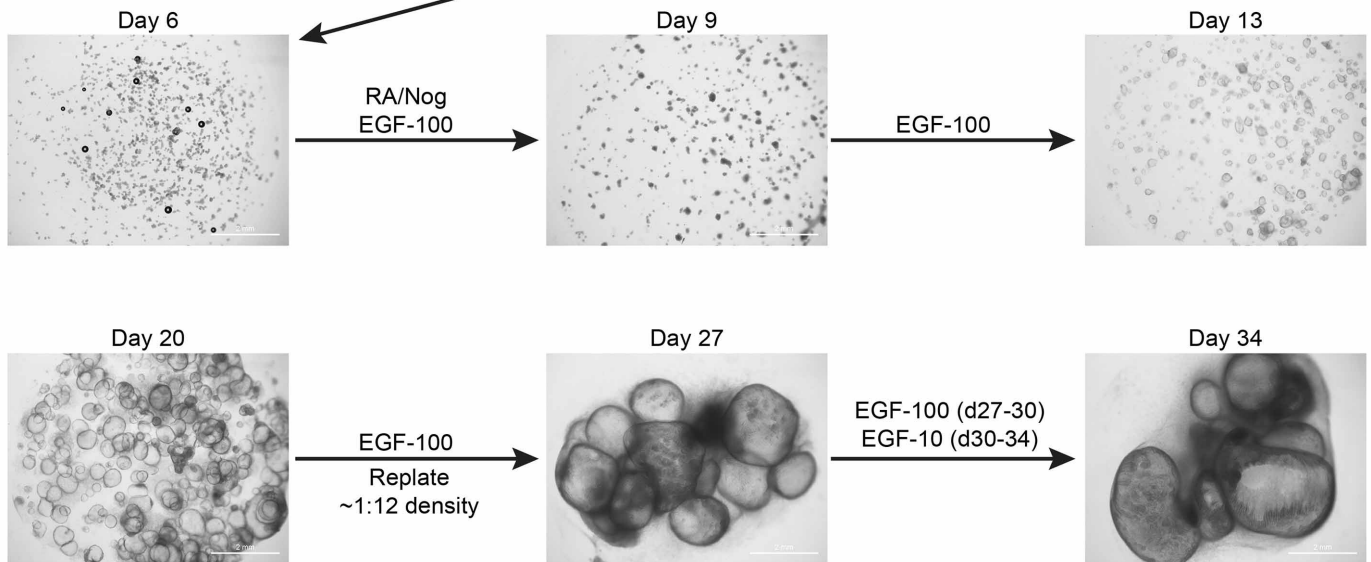
Extended Data Figure 9 | *H. pylori* infection of hGOs. **a**, hGOs were used to model human-specific disease processes of *H. pylori* infection. Bacteria were microinjected into the lumen of hGOs and bacteria were detected in the lumen 24 h after injection by bright-field microscopy (black arrow). **b**, Electron micrograph illustrating the attachment of an *H. pylori* bacterium to an hGO

epithelial cell 24 h after injection. Scale bar, 500 nm. **c**, Western blots from Fig. 4 that show the molecular mass markers in the first lane. The darker exposure for the CagA western blot (CagA dark) was included to show the molecular mass markers (170 and 130 kDa).

Monolayer Culture



3-D Culture



Extended Data Figure 10 | Summary of methods for the directed differentiation of gastric organoids. Each step in the differentiation process is indicated, along with representative stereomicrographs. EGF-100 and EGF-10 represent EGF at 100 ng ml^{-1} and 10 ng ml^{-1} , respectively.

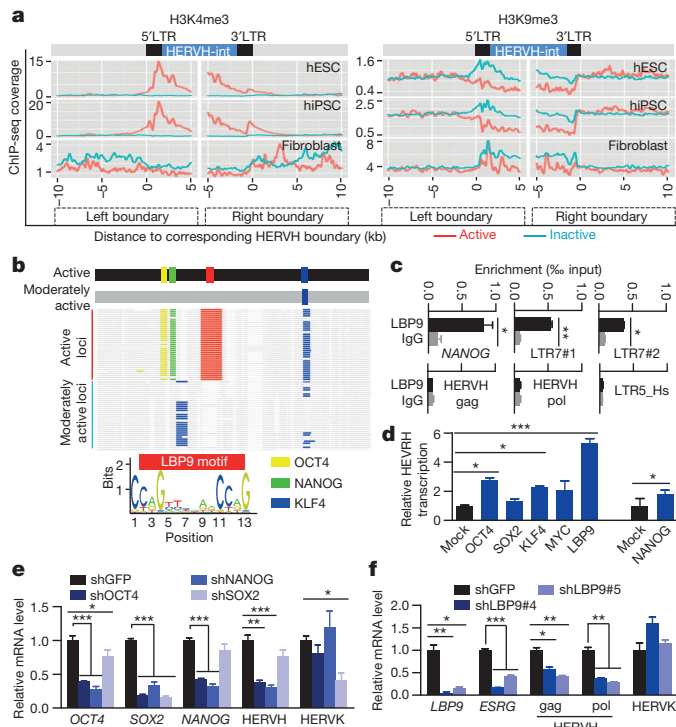


Figure 2 | HERVH is recruited into the circuitry of human pluripotency. **a**, The distribution of H3K4me3 and H3K9me3 in active versus inactive HERVH regions in hiPSCs, hESCs and HFF-1. **b**, Conserved binding sites of OCT4, NANOG, LBP9 and KLF4 are shown in active LTRs versus moderately active versions of LTR7C/Y. The Jasper consensus sequence of the LBP9 motif is shown. **c**, Confirmation of LBP9 binding to LTR7 by ChIP-qPCR with two different primers (LTR7#1 and LTR7#2) targeting LTR7 regions. HERVH gag, HERVH pol and LTR5_Hs (LTR of HERVK) serve as negative controls, while an upstream region of NANOG (7.5 kb from TSS) is a positive control. Data are collected from two independent experiments with biological replicates per experiment (LBP9, $n = 3$; IgG, $n = 2$), error bars indicate s.d.; t -test $*P < 0.05$, $**P < 0.01$. **d**, Upregulation of HERVH transcription in HFF-1 regulated by exogenous pluripotency-associated transcription factors. Data are collected from three independent experiments with biological triplicates per experiment. **e**, **f**, Effects of shRNA knockdowns of various transcription factors on HERVH and HERVK transcription in hESC_H9. Data shown are representative of three independent experiments with biological triplicates per experiment. **d–f**, Error bars indicate s.d.; t -test $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

(Extended Data Fig. 3a). A candidate KLF4 binding site was also identified within the LTR of HERVH (Fig. 2b). We additionally asked which transcription factor motifs are significantly enriched across four *in silico* tests (Extended Data Fig. 3b). Only one—LTR-binding protein 9 (LBP9) (also called murine Tfcp2l1)—was significant across all analyses (Extended Data Fig. 3b). Tfcp2l1 is within the Oct4 interactome¹⁶ and binds regulatory regions of *Oct4* and *Nanog*¹⁷ in mouse embryonic stem cells (mESCs). LBP9's direct binding to LTR7 is confirmed by ChIP-qPCR and electrophoretic mobility shift assay (EMSA) (Fig. 2c and Extended Data Fig. 3c). EMSA further demonstrates LBP9 and NANOG cooperation in binding LTR7 (Extended Data Fig. 3c), consistent with synergy after simultaneous overexpression (Extended Data Fig. 7c). LBP9-specific binding was also detected in the 5' region of NANOG (Fig. 2c).

In vitro differentiation assays showed that HERVH transcription levels declined over time in parallel with declines in *OCT4*, *NANOG* and *LBP9* (Extended Data Fig. 3d), suggesting a role in HERVH expression. As expected, ectopic expression of *LBP9*, *OCT4*, *NANOG* and *KLF4* activated the pT2-LTR7-GFP#2 reporter and enhanced endogenous HERVH transcription levels in human primary fibroblast (HFF-1), whereas overexpression of *MYC* or *SOX2* had no effect (Fig. 2d and Extended Data Fig. 7c). Conversely, a complementary 'loss of function' RNA interference (RNAi) assay in hESC_H9 revealed that HERVH transcription levels

were reduced after knockdown of *OCT4*, *NANOG* and *LBP9*, but not *SOX2* (Fig. 2e, f).

We confirmed that LBP9 directly stimulates HERVH-driven expression, by comparing signals of a wild-type pT2-LTR7-GFP#1 reporter construct and a mutant lacking the LBP9 motif (Δ LBP9) in hiPSCs (Extended Data Fig. 7d). When wild-type and mutant constructs were transfected into hiPSCs, the GFP signal was clearly detected from the wild-type reporter, but it was decreased by twofold in Δ LBP9 (Extended Data Fig. 7d).

Embryonic stem-cell-specific transcription factors *OCT4*, *NANOG*, *KLF4* and *LBP9* thus drive HERVH transcription in hPSCs. In contrast to mice in which LBP9 binding sites are genomically distinct from those of other pluripotency transcription factors⁶, the key pluripotent transcription factors cluster within the primate-specific HERVH (Fig. 2b).

To test the functional importance of HERVH, we analysed RNA-seq data to investigate the influence of HERVH on the expression of neighbouring regions. We find that LTR7 of HERVH initiates chimaeric transcripts, functions as an alternative promoter or modulates RNA processing from a distance (Fig. 3a, Extended Data Fig. 4b and Supplementary Tables 8 and 9). A total of 128 and 145 chimaeric transcripts were identified in hiPSCs and hESCs, respectively (Extended Data Fig. 4a and Supplementary Tables 8 and 9). One gene can contribute to multiple chimaeric transcripts. The chimaeric transcripts between HERVH and a downstream gene generally lack the 5' exon(s) of the canonical version (for example, *SCGB3A2*) while part of HERVH is exonized (for example, *RPL39L*) (Fig. 3a). A significant fraction of HERVH sequence can be incorporated into novel, lineage-specific genes (for example, *ESRG*; Fig. 3a) or long non-coding RNAs (lncRNAs) (for example, RP11-69I8.2; Extended Data Fig. 4d and Supplementary Table 10). We confirmed several hPSC-specific chimaeric transcripts by RT-PCR (Fig. 3a). Transcriptional start signals commonly map to 5' LTR-HERVH boundary regions (Extended Data Fig. 4c). Unlike the chimaeric transcripts, the canonical genes are commonly not expressed in pluripotent cells.

Nearly 10% of the transcripts driven by HERVH are annotated as lncRNA¹² (see Supplementary Table 11 for coding potential). Fifty-four transcripts were identified that are commonly detected in hPSCs, while the rest showed sporadic detection (Extended Data Fig. 4d). The former set includes linc-ROR and linc00458, known to modulate pluripotency^{18,19}. Alignment of the 22 most highly expressed transcripts reveals an LTR-HERVH-derived conserved core domain (CD) (Extended Data Fig. 4f). The domain is predicted to bind RNA-binding proteins, including pluripotency factors (for example, *NANOG*) and pluripotency-associated histone modifiers (for example, *SET1A* and *SETDB1*) (Extended Data Fig. 4g). In agreement with a role in pluripotency, linc00458 physically interacts with *SOX2* (ref. 19).

To explore the effect of either LBP9 or specific HERVH-derived transcripts on the reprogramming process, we investigated whether forced expression of *LBP9*, *ESRG* or the conserved domain of lncRNAs (LTR7-CD) modulates the fibroblast–hiPSC transition. While the overexpressed gene products affect neither pluripotency nor self-renewal (Extended Data Fig. 5a, b), all facilitate reprogramming by accelerating the mesenchymal–epithelium transition or hiPSC maturation (Fig. 3b and Extended Data Fig. 5c).

While LBP9 is key to the murine naive state^{6,20}, HERVH is primate-specific. To determine whether HERVH–LBP9 delineates a primate-specific pluripotency circuitry, we performed 'loss of function' experiments using small hairpin RNAs (shRNAs) against *LBP9* or HERVH (Fig. 3c–f and Extended Data Fig. 5d–g). Pluripotency-associated transcription factors and markers are downregulated whereas multi-lineage differentiation markers are upregulated upon knockdown of either *LBP9* or HERVH, but not in shRNA controls (shGFP) (Fig. 3c, d and Extended Data Fig. 5f, g). Depletion of *LBP9* or HERVH in hESCs thus results in loss of self-renewal. Knockout of *LBP9* similarly abolishes hESC self-renewal (Extended Data Fig. 5h–j). In contrast to hPSCs, the *Tfcp2l1* (*LBP9*) knockdown in mESCs does not reduce levels of *Oct4*, *Sox2* and *Nanog* in serum-based conditions (Extended Data Fig. 5k)²¹, but only

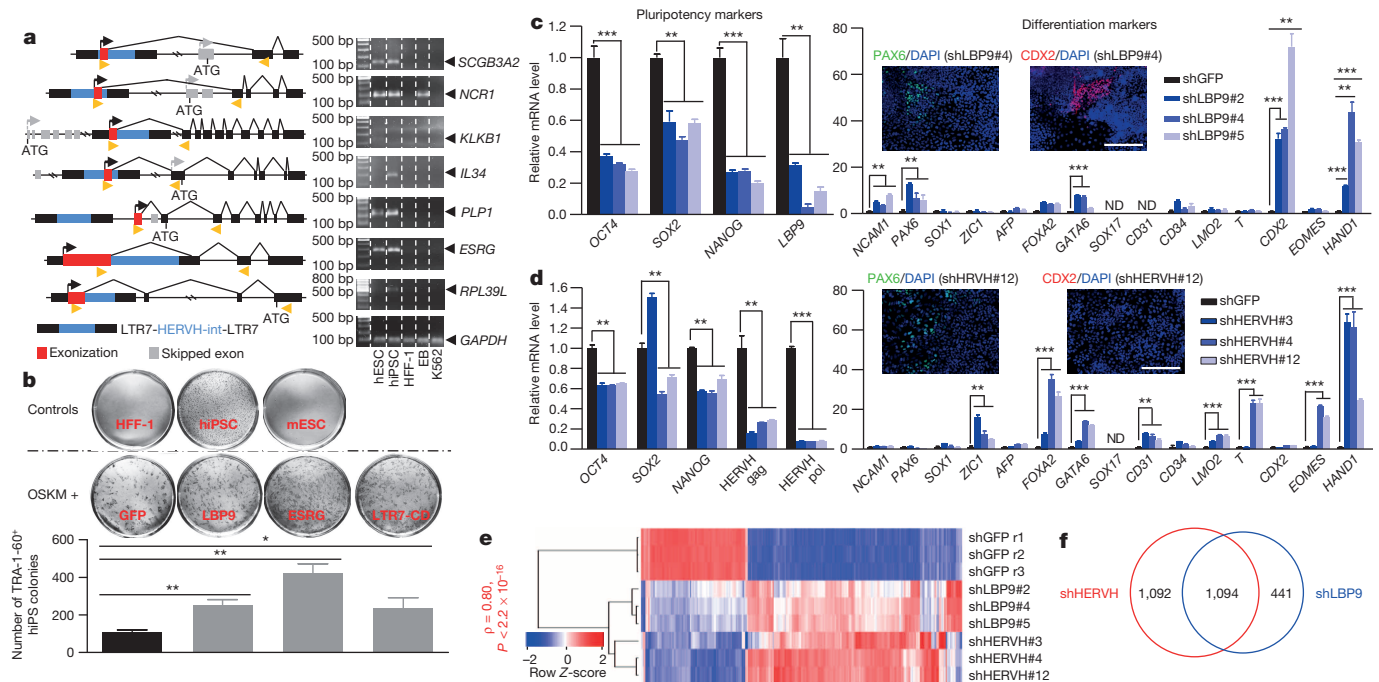


Figure 3 | HERVH triggers pluripotency-regulating hPSC-specific chimaeric transcripts and lncRNAs. **a**, Expression of HERVH forces diversification of transcripts in hPSCs. Left: schematic representation of the HERVH-derived alternative and chimaeric transcripts. Right: RT-PCR detects HERVH-specific transcripts (marked by triangles) in hPSCs (HERVH-NCRI also detected in embryoid body (EB)), but not in HFF-1 or K562 (myelogenous leukemia cell line). Yellow arrows indicate primer binding sites. **b**, The effects of *LBP9* and HERVH-derived transcripts on reprogramming of HFF-1 to hiPSCs. Top: representative TRA-1-60-stained wells are shown. Bottom: the number of TRA-1-60⁺ hiPS colonies reprogrammed from HFF-1 by *LBP9*, *ESRG* or LTR7-CD in conjunction with *OCT4*, *SOX2*, *KLF4* and *MYC* (OSKM). Error bars indicate s.d., *t*-test **P* < 0.05, ***P* < 0.01 from three independent experiments. **c**, **d**, qRT-PCR analyses to determine the relative expression level

of pluripotency and differentiation markers after knockdown of *LBP9* (**c**) or HERVH (**d**) in hESC_H9. Data shown are representative of three independent experiments with biological triplicates per experiment. Error bars indicate s.d., *t*-test **P* < 0.05, ***P* < 0.01, ****P* < 0.001. ND, not detected. Representative immunostainings show the expression of PAX6 and CDX2 in *LBP9* and HERVH knockdowns (scale bar, 200 μm). **e**, Heat map showing genome-wide gene expression in hESC_H9 after knockdown of GFP (shGFP), *LBP9* (shLBP9) and HERVH (shHERVH). The knockdown effect of *LBP9* and HERVH are highly similar (ρ from Spearman's correlation). For a list of affected genes, including direct targets of shHERVH, see Supplementary Tables 13 and 14. **f**, Venn diagram shows that 1,094 out of 2,627 genes are similarly affected by HERVH knockdown and *LBP9* knockdown (Supplementary Table 12).

in 2i (inhibition of ERK1 and GSK3 β signalling with small molecules)⁶. In fact, Tfc2l1 does not affect self-renewal, but rather differentiation potential (Extended Data Fig. 5k).

Genome-wide gene expression patterns are highly similar between *LBP9* and HERVH knockdowns (Fig. 3e), consistent with *LBP9* regulating HERVH-driven expression. A total of 1,094 of the 2,627 genes are similarly regulated in *LBP9* and HERVH knockdowns (Fig. 3f and Supplementary Table 12). While some HERVH-derived chimaeric transcripts are potentially directly affected by depletion of HERVH (Supplementary Tables 13, 14), qRT-PCR identified 19 HERVH-derived lncRNAs downregulated in response to both HERVH and *LBP9* knockdowns (Extended Data Fig. 4e).

While several of the differentially expressed genes are associated with murine pluripotency, the *LBP9*-HERVH-driven list of transcripts defines a primate-specific pluripotency network. Our analyses defined two classes of genes. Class I genes contained those conserved between mouse and human that contribute to the pluripotency in both; class II contained a primate-specific group that includes (a) those with an orthologous partner, but are not involved in murine pluripotency, and (b) novel (not in mouse) transcripts (Extended Data Fig. 4b, d). Several HERVH elements in class IIa affect gene expression in *cis*, and drive specific genic isoforms (for example, *SCGB3A2*). A subset of class IIb contains HERVH-derived novel sequences (for example, linc-ROR, linc000548, *ESRG*) (Extended Data Fig. 4d).

We examined one class IIb transcript in detail. *ESRG* has a putative open reading frame (ORF) only in human (Extended Data Fig. 6a and Supplementary Data 1), and is uniquely expressed in human inner cell mass (ICM) and PSCs (Extended Data Fig. 6b). Knockdown of *ESRG*

compromised self-renewal of hESCs, as expression of many pluripotency-associated genes was decreased, while *SOX2* expression was slightly elevated (Extended Data Fig. 6c–e). The *ESRG* knockdown colonies lost their hESC morphologies and committed to differentiation (Extended Data Fig. 6e, f). Expression of *ESRG* along with *OCT4*, *SOX2*, *KLF4* and *MYC* (OSKM) has a similar effect on the reprogramming process compared with *LBP9* (Extended Data Fig. 5c). *ESRG* is thus an HERVH-associated novel gene required for human-specific pluripotency, with a more specific phenotype than upstream regulators.

Given that the naive-associated transcription factors together cluster on HERVH and the HERVH-derived products are essential for primate pluripotency, we investigated whether HERVH-driven transcription marks the naive-like state in hPSC cultures. To explore this the reporter construct pT2-LTR7-GFP#2 was integrated into the genome of either mouse or human PSCs (Fig. 4a and Extended Data Figs 7a, b and 8i) by Sleeping Beauty gene transfer, providing stable transgene expression²². While all of the mESC colonies homogeneously express GFP (Extended Data Fig. 7a), only ~4% of cells in each hESC colony show a strong GFP signal (GFP^{high}), indicating cellular heterogeneity (Extended Data Fig. 7e, h–j). The fraction either weakly or not expressing GFP we term GFP^{low} and GFP[−], respectively (Fig. 4a and Extended Data Fig. 7b, e). RNA-seq data of hESCs from single cells^{23,24} and hPSC lines confirm that pluripotent cultures exhibit variability in HERVH expression (Extended Data Fig. 1d), indicating that the GFP^{high} subpopulation may differ from the GFP^{low} subpopulations. Consistent with a naive-like state, data mining of single-cell RNA-seq data sets²⁴ revealed that the expression level of HERVH in hESCs is correlated with several pluripotency-associated genes, including naive-associated transcription factors (Extended Data Fig. 1e).

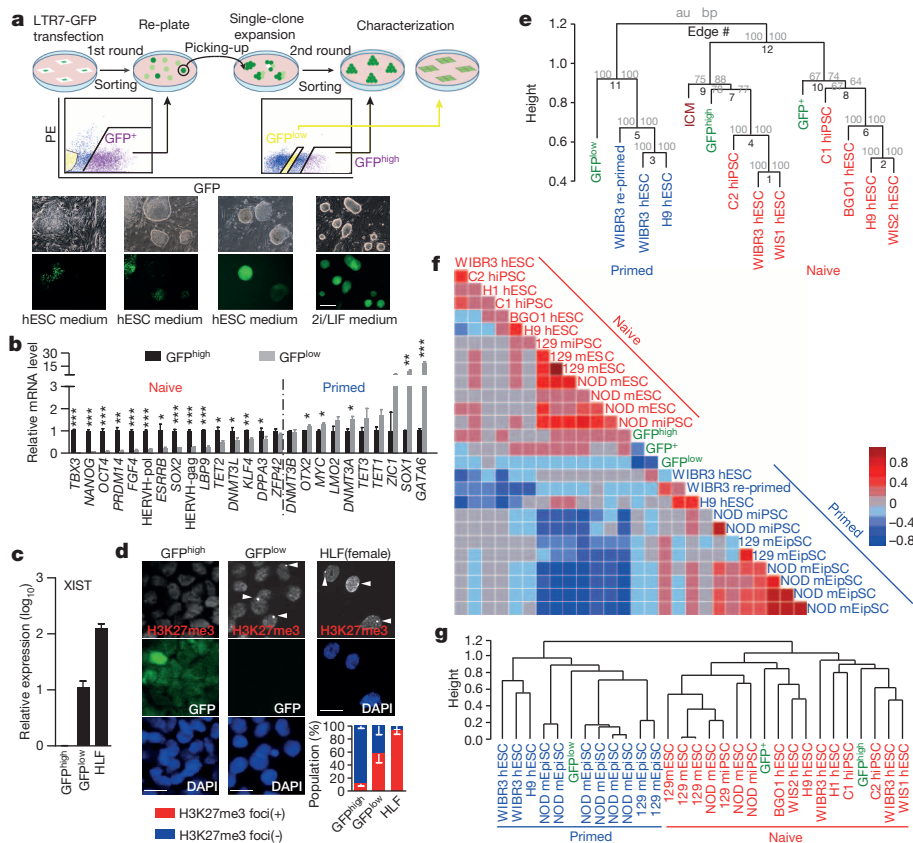


Figure 4 | HERVH genetically marks naive-like hESCs. **a**, Experimental scheme for isolating naive-like hPSCs. pT2-LTR7-GFP#2-marked hESC_H9 were enriched by FACS sorting in multiple rounds and cultured in conventional hESC medium and in 2i/LIF medium, respectively. Scale bar, 200 μ m. See also Supplementary Videos 1 and 2. **b**, qRT-PCR analyses of multiple transcription factors and markers for naive and primed state in GFP^{high} and GFP^{low} cells, respectively. **c**, qRT-PCR analysis of *XIST* in GFP^{high} and GFP^{low} hESC_H9 and human female fibroblasts (HLF). **b**, **c**, Error bars indicate s.d.; t-test * P < 0.05, ** P < 0.01 and *** P < 0.001 (n = 3 independent cell cultures). **d**, Representative confocal images obtained after immunostaining for H3K27me3 on GFP^{high} and GFP^{low} hESC_H9 and HLF. Scale bar, 20 μ m. The proportions of cells with (+) and without (−) H3K27me3 foci (arrowheads) in each sample are shown in the histogram.

To collect uniform GFP^{high} and GFP^{low} hPSCs, we performed two rounds of fluorescence-activated cell sorting (FACS) (Fig. 4a). We first sorted GFP⁺ cells that were further divided into GFP^{high} and GFP^{low} categories. Notably, GFP^{high} cells are capable of forming tight, uniformly expressing three-dimensional (3D) colonies characteristic of naive mESCs (Fig. 4a and Supplementary Video 1). In contrast, GFP^{low} cells form flat colonies, resembling mouse epiblast stem cells (mEpiSCs) (Fig. 4a). We also observed mosaic colonies. Immunostaining of 3D and chimaeric colonies revealed that the NANOG and GFP^{high} signals co-present (Supplementary Videos 1 and 2). Thus, the GFP^{high} subpopulation in human pluripotent stem cells is enriched for cells resembling the murine naive/ground state.

To examine this possibility, we subjected GFP^{high} and GFP^{low} cells to expression analyses. qRT-PCR revealed significant upregulation of naive-associated transcription factors^{4–6} and downregulation of lineage-commitment genes in GFP^{high} versus GFP^{low} cells (Fig. 4b). As in naive mESCs²⁵ and human ICM²⁶, X chromosomes are activated in GFP^{high} hESC_H9, as evidenced by nearly complete loss of condensed H3K27me3 nuclear foci (Fig. 4d) and low levels of *XIST* expression (Fig. 4c). However, nearly 60% of GFP^{low} hESCs that transitioned from GFP^{high} hESCs are marked with condensed H3K27me3 foci or higher density of H3K27me3 in the nucleus (Fig. 4d and Extended Data Fig. 8g). These data are consistent with a naive-like state for GFP^{high} cells and a primed state for

GFP^{low} cells (one X chromosome inactivated or in the process of being inactivated). GFP^{high} cells can be maintained in the modified 2i/LIF medium for a long time, with higher single-cell clonality as well as full pluripotency (Extended Data Fig. 8a–d). However, GFP^{high} and GFP^{low} cells have slightly different differentiation potential. When differentiation is triggered, certain naive-associated transcription factors are maintained at higher levels in GFP^{high} naive-like cells compared with GFP^{low} cells, and start their differentiation program with a delay (Extended Data Fig. 8e, f). Early passage hPSC cultures behave similarly to GFP^{high} cells (Extended Data Fig. 9a–c).

Transcriptomes of GFP-sorted cell populations and previously characterized naive-like and primed hPSCs⁴ and mouse counterparts, as well as human ICM, support a naive-like status of GFP^{high} cells. Unbiased hierarchical clustering of the expression profiles revealed that GFP^{high} and GFP⁺ cells have a similar, but non-identical, expression pattern, one that sharply contrasts with the expression pattern of GFP^{low} cells (Extended Data Fig. 8h). Notably, GFP^{high} and GFP⁺ samples clustered with human ICM and the published naive-like hPSCs, respectively (Fig. 4e). Importantly, GFP^{high} cells clustered closest to human ICM (Fig. 4e).

Cross-species comparison of expression of 9,583 mouse–human orthologues revealed that GFP^{high} and GFP⁺ correlated to published naive cells, while GFP^{low} clustered with primed cells (Figs 4f, g), supporting

the significance of HERVH-driven transcription defining a naive-like state.

To address how gene expression changes up to the ICM stage, we analysed 114 RNA-seq samples harvested in early developmental stages of embryogenesis²⁴ and 3 RNA-seq samples of naive-like hESCs (3iL_hESC³). HERVH expression appears already in the zygote, but the pattern of activated loci changes during early development (Extended Data Fig. 9d, e). Importantly, the pattern of active loci characteristic of ICM is the closest to naive-like hESCs, including GFP^{high} cells (Extended Data Fig. 9d). Notably, consistent with the active chromatin state, the number of activated HERVH loci in naive-like cells is particularly higher than in the primed hESCs (Extended Data Fig. 9d–f), indicating that HERVH may have some involvement in the derivation and/or maintenance of naive-like hPSCs.

To address how HERVH-driven gene expression modulates pluripotency, we surveyed differentially regulated genes in GFP^{high} versus GFP^{low} cells, intersected by HERVH cis-regulation. The differentially regulated genes located in the neighbourhood (± 50 kb) of HERVH display a similar expression pattern to those differentially expressed in GFP^{high} versus GFP^{low} and in human naive-like versus primed stages, derived under specific culture conditions⁴ (Extended Data Fig. 9h). In contrast, a distinct pattern is observed when comparing mESCs versus mEpiSCs (Extended Data Fig. 9g). Strikingly, there is an inverse pattern of expression between genes defining a naive-like stage (upregulated in GFP^{high} versus GFP^{low}) and those that are downregulated in HERVH knockdowns ($\rho = -0.6$, $P \ll 0.0001$; Extended Data Fig. 9i), underlying the significance of HERVH in regulating the naive-like state in humans. Differentially expressed genes between GFP^{high} versus GFP^{low} populations were enriched for Gene Ontology (GO) terms of developmental processes, morphogenesis and organismal processes (Extended Data Fig. 9j). Transition of naive-like cells into a primed state after depletion of HERVH supports the above conclusion (Extended Data Fig. 9k).

While GFP^{high} cells have many properties resembling naive mESCs, they are better regarded as being naive-like, not least because it is unclear that human and naive mESCs need be identical. Indeed, while LBP9 is associated with pluripotency^{6,20} in mammals, HERVH was recruited to the pluripotency network exclusively in primates. How then to define naive human pluripotency if we do not necessarily expect cells in this state to be identical to those in mouse? We suggest that, rather than hard-to-replicate inter-species chimera experiments²⁷, the optimal approach is to define cells by similarity of expression to the ICM (see Supplementary Discussion). In this regard, GFP^{high} cells are one of the best current models of naive-like status.

That LBP9 forms heteromer complexes functioning either as a transcriptional activator or a repressor, depending on the partner²⁸, is consistent with HERVH being recruited to the pluripotency network by serendipitous modification of a pluripotency factor detailed to defend the cell against it (Extended Data Fig. 10). Whatever the origin, LTR7 of HERVH is an efficient reporter for the naive-like state most probably because it acts as a platform for multiple key pluripotent transcription factors²⁹. Similarly, the LTR7–GFP reporter should enable optimization of naive-like hPSC culture conditions.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 March; accepted 28 August 2014.

Published online 15 October; corrected online 17 December 2014 (see full-text HTML version for details).

1. Welling, M. & Geijsen, N. Uncovering the true identity of naive pluripotent stem cells. *Trends Cell Biol.* **23**, 442–448 (2013).
2. Ware, C. B. *et al.* Derivation of naive human embryonic stem cells. *Proc. Natl Acad. Sci.* **111**, 4484–4489 (2014).
3. Chan, Y. S. *et al.* Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell* **13**, 663–675 (2013).
4. Gafni, O. *et al.* Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282–286 (2013).

5. Hanna, J. *et al.* Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc. Natl Acad. Sci. USA* **107**, 9222–9227 (2010).
6. Martello, G., Bertone, P. & Smith, A. Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. *EMBO J.* **32**, 2561–2574 (2013).
7. Kunarso, G. *et al.* Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nature Genet.* **42**, 631–634 (2010).
8. Lu, X. *et al.* The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nature Struct. Mol. Biol.* **21**, 423–425 (2014).
9. Fort, A. *et al.* Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance. *Nature Genet.* **46**, 558–566 (2014).
10. Macfarlan, T. S. *et al.* Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).
11. Santoni, F. A., Guerra, J. & Luban, J. HERV-H RNA is abundant in human embryonic stem cells and a precise marker for pluripotency. *Retrovirology* **9**, 111 (2012).
12. Kelley, D. & Rinn, J. Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol.* **13**, R107 (2012).
13. Gaspar-Maia, A. *et al.* Chd1 regulates open chromatin and pluripotency of embryonic stem cells. *Nature* **460**, 863–868 (2009).
14. Chappell, J., Sun, Y., Singh, A. & Dalton, S. MYC/MAX control ERK signaling and pluripotency by regulation of dual-specificity phosphatases 2 and 7. *Genes Dev.* **27**, 725–733 (2013).
15. Xie, W. *et al.* Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* **153**, 1134–1148 (2013).
16. van den Berg, D. L. *et al.* An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell* **6**, 369–381 (2010).
17. Chen, X. *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106–1117 (2008).
18. Loewer, S. *et al.* Large intergenic non-coding RNA-ROR modulates reprogramming of human induced pluripotent stem cells. *Nature Genet.* **42**, 1113–1117 (2010).
19. Ng, S. Y., Johnson, R. & Stanton, L. W. Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. *EMBO J.* **31**, 522–533 (2012).
20. Ye, S., Li, P., Tong, C. & Ying, Q. L. Embryonic stem cell self-renewal pathways converge on the transcription factor Tfc2l1. *EMBO J.* **32**, 2548–2560 (2013).
21. Nishiyama, A. *et al.* Systematic repression of transcription factors reveals limited patterns of gene expression changes in ES cells. *Sci. Rep.* **3**, 1390 (2013).
22. Mátés, L. *et al.* Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nature Genet.* **41**, 753–761 (2009).
23. Ramsköld, D. *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nature Biotechnol.* **30**, 777–782 (2012).
24. Yan, L. Y. *et al.* Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nature Struct. Mol. Biol.* **20**, 1131–1139 (2013).
25. Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487–492 (2009).
26. Okamoto, I. *et al.* Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature* **472**, 370–374 (2011).
27. Theunissen, T. W. *et al.* Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* <http://dx.doi.org/10.1016/j.stem.2014.07.002> (2014).
28. To, S., Rodda, S. J., Rathjen, P. D. & Keough, R. A. Modulation of CP2 family transcriptional activity by CRT1-1 and sumoylation. *PLoS ONE* **5**, e11702 (2010).
29. Dunn, S. J., Martello, G., Yordanov, B., Emmott, S. & Smith, A. G. Defining an essential transcription factor program for naive pluripotency. *Science* **344**, 1156–1160 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements L.D.H. is Wolfson Royal Society Research Merit Award Holder. A.T.G. is funded by a scholarship from the University of Bath. Z.Iz. is funded by ERC-2011-AdG 294742. G.G.S. is funded by DFG grant SCHU1014/8-1 and LOEWE Center for Cell and Gene Therapy Frankfurt/Hessian Ministry of Higher Education, Research and the Arts (ref. number III L 4-518/17.004). We thank U. Martin and S. Merkert (Leibniz Research Laboratories for Biotechnology and Artificial Organs (LEBAO), Hannover Medical School, Hannover, Germany) for providing the cell lines hCBEC, hCBIP51, hCBIP52 and hFF-IP54. We thank G. Klein for the inspiration of working with ERVs and Z. Cserenyés for his assistance in imaging.

Author Contributions This project was inspired by M.C.L. Z.Iz., L.D.H. and J.W. conceived ideas for the project, and wrote the manuscript with contributions from other authors. The project was supervised by Z.Iz. and L.D.H. Z.Iz. provided critical advice. J.W. designed and performed experiments, analysed and interpreted data, and participated in bioinformatic analyses. T.R. contributed by EMSA and assisted in immunostaining experiments. A.S. assisted in the reporter assays. H.C. assisted in shRNA cloning. W.C. and J.W. performed RNA-seq experiments. A.P. provided materials and performed karyotype analysis. D.B., N.V.F. and G.G.S. provided materials. G.X. performed RNA-seq, bisulfite-seq and ChIP-seq analyses. M.S. analysed microarray data and performed cross-species correlation studies. L.D.H. and A.T.G. performed all the other bioinformatic analyses.

Author Information RNA-seq and microarray data were submitted to NCBI's GEO database under accession GSE54726. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Z.Iz. (zizvak@mdc-berlin.de) or L.D.H. (l.d.hurst@bath.ac.uk).

METHODS

Ethics approval. For work on human ES cells we obtained No. 6 allowance from the Robert Koch Institute, Germany (8 October 2004). The human embryonic stem cell lines (H1, H9, BGN1 and BGN2) are permitted to be used in the study “Mechanisms of single transduction in the maintenance of undifferentiated state in human embryonic stem cells”.

Cell culture. Human foreskin fibroblasts (HFF-1) (ATCC, SCRC-1041) were cultured with the fibroblast medium (DMEM, 20% FBS, 1 mM L-glutamine, 1% non-essential amino acids, 0.1 mM 2-mercaptoethanol and primocin), and were passaged every 3–4 days. Human embryonic stem cells (hESCs) were cultured in matrigel/feeder-coated plates in the conventional hESC medium (knockout DMEM, 20% knockout serum supplement, 1 mM L-glutamine, 1% nonessential amino acids, 0.1 mM 2-mercaptoethanol, 10 ng ml⁻¹ bFGF (Pepro Tech, 100-18B) and primocin), or in naive hESC media NHSM⁴ or 3iL³ or in human 2i/LIF medium (this work). The human 2i/LIF medium is based on mouse 2i/LIF medium⁶ (knockout DMEM, 20% knockout serum replacement, 1 mM L-glutamine, 1% nonessential amino acids, 0.1 mM 2-mercaptoethanol, 10 ng ml⁻¹ LIF, 3 μM CHIR99021, 1 μM PD0325901 and primocin, but the CHIR99021 was changed from 3 to 1 μM, and the medium was supplemented with 10 ng ml⁻¹ bFGF). The medium was changed daily. hESCs were treated with collagenase IV (1 mg ml⁻¹) (Life Technologies, 17104-019) and then passaged onto new matrigel/feeder-coated plates every 4–5 days. The generation of hiPSC line hiPS-SB4 and hiPS-SB5 has been reported³⁰. iPSC lines hCBiPS1 and hCBiPS2 and their culture conditions have been described previously³¹. They were derived from human cord-blood-derived endothelial cells (hCBEC) using a lentiviral vector expressing reprogramming factors *OCT4*, *SOX2*, *NANOG* and *LIN28* (ref. 31). Similarly, the line hFF-iPS4 (previously known as hiPS-SK4) was produced using HFF-1 cells and the same lentiviral overexpression construct. Successful reprogramming for the hFF-iPS4 cell line was verified by morphology, the expression of pluripotency markers, karyogram analysis and the ability to generate teratomas on immunocompromised mice (data not shown).

Mouse ESCs were cultured in gelatin/feeder-coated plates with the mESC medium (knockout DMEM, 15% fetal calf serum (FCS), 1 mM L-glutamine, 1% nonessential amino acids, 0.1 mM 2-mercaptoethanol, 10 ng ml⁻¹ LIF (Millipore, LIF1010) and primocin) or mouse 2i/LIF medium⁶. To prepare feeders, mouse embryonic fibroblasts (passage 4) isolated from CF-1 mouse embryos were treated with mitomycin C (10 μg ml⁻¹) for 2–3 h.

All above-mentioned cell cultures tested negative for mycoplasma infection. Karyotype of hESC_H9 was analysed using the G-banding method³² indicating normal karyotype (Extended Data Fig. 8j).

Reprogramming assay. Reprogramming was performed as described previously^{30,33}. Briefly, 200,000 HFF-1 cells were transfected with pT2/RMCE-OSKM (2 μg) and pT2-CAG-amaxaGFP, or pT2-CAG-HA-LBP9, or pT2-CAG-ESRG, or pT2-LTR7-CD (1 μg per plasmid) using the Neon transfection system (Life technologies), and transposition was induced by SB100X²² (1 μg). The transfected cells were plated onto matrigel-coated 6-well plates and cultured in the fibroblast medium (first 2 days), then medium was changed to the hESC medium (day 2 post-transfection). After 3 weeks, several hESC-like colonies were picked for expansion and characterization, while the rest of the colonies were fixed in 4% with paraformaldehyde and subjected to immunostaining.

In vitro differentiation assay. To spontaneously differentiate hPSCs to embryoid bodies (EBs), hESCs/hiPSCs were cultured on geltrex-coated 6-well plates. Cells from one well were dissociated with collagenase IV (1 mg ml⁻¹) for 5 min, and then split into small cell clumps. The small cell clumps were transferred into three 10-cm low-attachment dishes, and cultured in EB medium (knockout DMEM, 20% knockout serum replacement, 1 mM L-glutamine, 1% nonessential amino acids, 0.1 mM 2-mercaptoethanol and primocin). The medium was changed every 2 days. The embryoid bodies were cultured for 10 days followed by collection for RNA-seq or then replated in gelatin-coated 6-well plates for one week followed by immunostaining.

Differentiation potential assay. GFP^{high} and GFP^{low} cells collected from the same FACS-sorted hESC clone were seeded on matrigel/feeder-coated plates, respectively. First, the GFP^{high} and GFP^{low} cells were cultured either in the human 2i/LIF medium or conventional hESC medium. Following 3 days culturing in the respective mediums, cells were exposed to EB medium. To improve single-cell-viability, the cells were treated with the ROCK inhibitor Y-27632 (Millipore, 10 μM) for 48 h before and after sorting.

Immunostaining. hPSC colonies were cultured on matrigel/feeder-coated chamber slides (BD Biosciences). Following 3 days of culturing, cells were fixed for 30 min in 4% paraformaldehyde, permeabilized for 30 min in 1% Triton X-100, and blocked for 1 h in blocking solution (Applied StemCell, ASB0103). Fixed cells were incubated overnight at 4 °C with primary antibodies (OCT4, SOX2, NANOG, SSEA4, TRA-1-60, PAX6, TUBB3 (βIII-tubulin), SOX17, α-SMA and CDX2) (Supplementary Table 3). After washing in PBS, the cells were incubated with secondary antibodies (Life Technology) for 1 h at room temperature. DAPI (Sigma, D9564) was

used for staining the nuclei. Immunostaining of reprogramming plates was performed as previously described³⁴. Briefly, cells were fixed with 4% paraformaldehyde and stained with biotin-anti-TRA-1-60 (eBioscience, 13-8863-80) and streptavidin horseradish peroxidase (Biolegend, 405210), diluted in 1% Triton X-100 (containing 0.3% BSA). Staining was performed using the Vector labs DAB kit (SK-4100). Stained hiPSC colonies were counted with ImageJ software.

Immunofluorescence microscopy to determine H3K27me3 in hESCs. GFP^{high} cells were seeded on matrigel-coated coverslips in 12-well culture plates. Following 4 days of culturing, the cells were fixed with 4% paraformaldehyde (Sigma) supplemented with DAPI for 15 min, and permeabilized with 0.5% Triton X-100 for 5 min. Fixed cells were incubated with primary antibodies (NANOG or H3K27me3, Novus Biologicals and Millipore, respectively) overnight at 4 °C, then washed three times with PBS, and incubated with secondary antibodies (Alexa Fluor, Life Technologies) for 1 hour. After additional washing, the samples were mounted using ProLong Gold antifade reagent (Invitrogen) and images were taken using a Zeiss LSM710 point-scanning single-photon confocal microscope. 3D image movies were created by Imaris Imaging Software (Bitplane). To statistically compare X chromosome state in GFP^{high} and GFP^{low} cells which were transited from GFP^{high}, images on GFP^{high} and GFP^{low} hESCs and female human fibroblast were analysed and quantified for the proportion of cells with condensed H3K27me3 foci which mark the inactive X chromosome. An average of 100–450 individual cells per samples from five images were counted.

DNA constructs. The *LBP9* ORF was amplified from human placenta cDNA by PCR with Pfu Ultra II Fusion HS (Agilent Technologies). A NotI restriction site was added to the 3' end of the fragment (for cloning purposes). A single, ~1,500-bp band was cloned into pJET1.2/blunt using the CloneJET PCR Cloning kit (Thermo Scientific). The *LBP9* fragment was re-amplified from pJET1.2-LBP9 plasmid digested with NotI and was cloned into pHA5 expression vector. The HA-LBP9 fragment was cut from pHA-CAG-HA-LBP9 vector and cloned into the Sleeping Beauty transposon³⁵, pT2-CAG-GFP vector. *LBP9* expression from pHA-CAG-LBP9 or pT2-CAG-HA-LBP9 was confirmed by western blotting. The size of the observed band was in good agreement with the molecular mass of the full-length protein (54,627 Da). *ESRG* was PCR amplified from hESC cDNA (Pfu Ultra II Fusion HS). The MluI and BglII restriction sites were added to the 5' and 3' ends, respectively, for subsequent cloning. A single ~300-bp band was digested with MluI and BglII restriction enzymes, and then cloned into pT2-CAG-GFP vector. To clone pT2-LTR7-CD, 22 highly expressed, HERVH-derived lncRNAs were first aligned (Clustal Omega alignment tool), and the lncRNA core domain (CD) sequence (Supplementary Table 1) was synthesized. The synthetic LTR7-CD flanked by MluI/BglII restriction sites was cloned into the pT2-CAG-GFP vector by replacing GFP.

Reporter assays. The individual HERVHs were compared with the HERVH consensus sequence from Repbase (<http://www.girinst.org/repbase/>). The *ESRG* locus of HERVH was selected to generate a reporter construct. Two different DNA fragments, #1 and #2, were amplified (for primers see Supplementary Table 1). LTR7#1 (566 bp) contains the *ESRG*-LTR7 flanked by ~110 bp upstream genomic sequence, while *ESRG*-LTR7#2 (1,194 bp) contains the LTR7 plus sequence from the HERVH-int. EcoRI and MluI restriction sites were added to the 5' and 3' ends of the fragments, respectively, for cloning purposes. The two DNA fragments were cloned into SB transposon-based pT2-CAG-GFP vector, digested with EcoRI and MluI (to remove CAG promoter) to generate pT2-LTR7-GFP#1 and pT2-LTR7-GFP#2. To clone an *LBP9*-motif deleted reporter construct, a 17-bp segment containing the *LBP9* motif was removed from pT2-LTR7-GFP#1 by inverse PCR (Extended Data Fig. 7d). The PCR-amplified ~5,600-bp fragment was gel-isolated (Qiagen), circularized and subsequently transformed into chemical competent DH5α cells. The deletion was confirmed by sequencing. The modified region was moved into the original vector by NcoI digestion. To generate multiple LTR7 reporter constructs (#3–#6), LTR7 was PCR-amplified from different genomic loci (Supplementary Table 1). The obtained fragments were gel isolated and cloned into pJET1.2 vector using the CloneJET PCR Cloning kit (Thermo Scientific) and confirmed by sequencing. In pT2-LTR7-GFP#3–#6, the LTR7 (flanked by StuI and Bsu36I) sequence of the pT2-LTR7-GFP#2 reporter was replaced by LTR7 (#3–#6). Finally, these vectors were transfected into fibroblasts and hiPSCs for subsequent analyses. The transfected fibroblasts and hiPSCs were cultured in the conventional hESC medium. GFP⁺ cells were quantified by FACS on day 6 post-transfection.

Gain-of-function assays. Individual expression plasmid constructs containing *OCT4*, *NANOG*, *SOX2*, *KLF4*, *MYC* or *LBP9* were transfected into 2 × 10⁵ HFF-1s, respectively. The transfected cells were collected for total RNA extraction and qRT-PCR on day 4 post-transfection.

Generating shRNA constructs. To generate shRNA against HERVH, we first aligned all active (based on RNA-seq data) full-length HERVHs and selected several conserved sequences. The selected conserved sequences were analysed by the Block-It RNAi Designer online program (<https://rnaidesigner.invitrogen.com/rnaexpress>). The shRNA sequences of score >3.5 were further analysed for their specificity using

BLAST against human genome. shESRG and shLBP9 targeting sequences were designed using the online siRNA design tool siDESIGN Center (<http://dharmacon.gelifsciences.com/design-center/>). 60-mer oligos were synthesized, and then cloned into the FP-H1 vector³⁶. shRNA targeting GFP was used as a control. NANOG, OCT4 and SOX2 shRNAs were previously described³⁷. Clones were verified by sequencing. For the list of shRNAs see Supplementary Table 2.

Generating stable shRNA knockdown hPSC lines. All of hESCs/hiPSCs were cultured under the same condition, including identical passage numbers. hESC/hiPSC cultures containing spontaneously differentiated cells (>10%) were excluded from the knockdown experiments. shRNA plasmid (10 µg) for each gene was transfected into 1×10^6 hPSCs by the Neon transfection system followed by G418 (500 µg ml⁻¹) selection on day 2 post-transfection until 7–10 days. Stable knockdown cell lines were harvested for FACS, immunostaining and RNA extraction.

Transfection of hPSCs. Cells were treated with ROCK inhibitor Y-27632 (10 µM) (Millipore, 688000) overnight before transfection, and then trypsinized with Accutase (Life Technologies, A1110501) for 3 min at 37 °C to generate single-cell suspension. 5×10^5 hiPSCs or hESCs were transfected with certain plasmids using the Neon transfection system. The transfected hPSCs were immediately re-plated onto the matrigel/feeder-coated 6-well plates in hESC medium containing Y-27632 (10 µM). Four hours post-transfection, the medium was refreshed to remove the transfection buffers and dead cells. The hESC medium was changed daily. Note that the Neon transfection system was also used to transfect HFF-1, mouse embryonic fibroblasts, and mESCs (according to the manufacturer's protocol).

Analysing hPSCs by FACS. Single-cell suspension was generated by treating hiPSCs/hESCs with Accutase for 3 min at 37 °C. 2×10^5 cells were incubated with anti-TRA-1-81-APC antibody (eBioscience, 17-8883-41) for 30 min at 4 °C in PBS. Cells were washed and suspended in ice-cold PBS before analysis on FACSCalibur (BD Biosciences). 10,000 cells were typically analysed.

Generating genetically LTR7-GFP marked hPSCs. Single-cell suspension of 5×10^5 hPSCs was transfected with 5 µg pT2-LTR7-GFP#2 and 500 ng SB100X using the Neon transfection system, and seeded onto matrigel/feeder-coated 6-well plates. One week post-transfection, hPSCs were treated with Y-27632 (10 µM) overnight, trypsinized into single cells, and purified with the feeder removal microbeads kit (Miltenyi Biotec, 130-095-531) before sorting by FACS. GFP⁺ and GFP⁻ were collected, respectively. The GFP⁺ hPSCs were re-plated on matrigel/feeder-coated 6-well plates and cultured in hESC medium. One week later, the single GFP⁺ colonies were picked up for expansion in hESC medium. The second round of sorting was performed on the expanded single clones to collect hPSCs expressing strong and low GFP signal (referred to as GFP^{high} and GFP^{low}), respectively. The GFP^{high} hPSCs were re-plated onto matrigel/feeder-coated 6-well plates and cultured in 2i/LIF medium for further characterization. The pT2-LTR7-GFP#2-marked individual hESC-H9 clones, GFP^{high}, GFP⁺ and GFP^{low}, were characterized in multiple assays. The integration site of the single copy pT2-LTR7-GFP#2 reporter in GFP^{high} was determined (Extended Data Fig. 8i).

Single-cell cloning assay. 1,000 GFP^{high} hESC_H9 collected from the second round of sorting were seeded onto one matrigel/feeder-coated well of the 6-well plate and cultured in 2i/LIF medium with or without Y-27632 (10 µM). 1,000 GFP^{low} hESC_H9s were seeded onto one matrigel/feeder-coated well of the 6-well plate and cultured in hESC medium with or without Y-27632 (10 µM). One week after seeding the hESCs were fixed with 4% paraformaldehyde for 1 min, and then stained with alkaline phosphatase (Sigma, ABO300). Pictures of stained cells were analysed. Dark blue (undifferentiated), light blue (partially differentiated) and colourless (differentiated) colonies were counted, respectively.

qRT-PCR. Total RNA was extracted from cells by using the Trizol kit (Invitrogen) following the manufacturer's instructions. 0.1 µg purified DNaseI-treated RNA, which was the mixture of biological triplicates, was used for reverse transcription (RT) (High Capacity RNA-to-cDNA kit, Applied Biosystems). Quantitative RT-PCR (qRT-PCR) was performed using the Power SYBR Green PCR Master Mix (Applied Biosystems) on the ABI7900HT sequence detector (Applied Biosystems). Data were normalized to GAPDH expression using the $\Delta\Delta C_t$ method. Error bars represent the standard deviation (s.d.) of samples carried out in triplicates. For the list of primers see Supplementary Table 1.

Electrophoretic mobility shift assay (EMSA). 2×10^6 hiPSCs were transfected with 20 µg plasmids encoding pT2-CAG-HA-LBP9. Two days post-transfection cells were collected and washed with PBS. Cells were lysed in 100 µl lysis buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10 mM EDTA, 5% glycerine, 1% NP-40 and 1× protease inhibitor cocktail (Roche)) for 30 min at 4 °C. Following removal of the cell debris by centrifugation at 20,000g, binding reactions were performed in 25 µl volumes at room temperature for 30 min. DNA binding reactions contained FAM-labelled LTR7-specific, complementary dsDNA oligonucleotides (LTR7 oligo), HA-LBP9 containing cell extracts, 10 mM Tris-HCl pH 8.5, poly(dI-dC), 1 mM EDTA, 50 mM KCl, 10 mM 2-mercaptoethanol (see also Extended Data Fig. 3c). Probe sequences are listed in Supplementary Table 1. The gel buffer contained 50 mM

Tris-borate pH 8.3, 1 mM EDTA. To super-shift specific complexes, cell extracts were incubated with antibodies (anti-LBP9 (Novus); anti-NANOG (Novus)) at 4 °C for 15 min before addition of the dsDNA oligonucleotides. Protein-DNA complexes were separated by electrophoresis in 6% non-denaturing polyacrylamide gels at 4 °C. Electrophoresis was performed at constant voltage of 200 V for 3, 4 or 6 h. The fluorescent signal was detected by using a FUJI FLA-3000 Imager.

ChIP-qPCR. ChIP-qPCR was performed with the Transcription ChIP kit (Diagenode) according to the manufacturer's instructions with slight modifications. 1×10^7 hPSCs were fixed in 1% formalin/hESC medium (v/v) for 10 min with gentle agitation on a rotator at room temperature. Fixation was stopped by the addition of glycine (125 mM) and agitation for 5 min at room temperature. Fixed cells were washed twice in ice-cold PBS, re-suspended in 15 ml lysis buffer. Chromatin was sheared by sonication to about 100–500-bp fragments using a Bioruptor (Diagenode) and diluted into immunoprecipitation buffer. Anti-LBP9 (Novus) and anti-IgG (Abcam) antibodies were added to sonicated chromatin solution and incubated with pre-blocked protein A magnetic beads (Invitrogen) overnight at 4 °C with gentle agitation on a rotator. Immune chromatin-bead precipitates were collected by the magnetic device (Invitrogen) at 4 °C. Precipitates were washed sequentially with washing buffer (Diagenode). Immunoprecipitated DNA was eluted by incubating the beads with 150 µl elution buffer with gentle agitation for 25 min at room temperature. To reverse crosslinking, sodium chloride (final concentration of 0.2 M) was added to the eluates that were incubated overnight at 65 °C. DNA was purified according to the manufacturer's instructions. Purified DNA from input and immunoprecipitation was used as templates for Taqman qPCR to determine the occupancy of LBP9 on NANOG, LTR7, HERVH-int (gag and pol) and LTR5_Hs. Primer and probe sequences are listed in Supplementary Table 1.

Analysis of genomic integration sites of the reporter construct in hESCs. The reporter LTR7-GFP#2 was cloned into Sleeping Beauty-based cloning vector pT2. The reporter was integrated into hESCs_H9 by co-transfecting the SB100X transposase²². Using sorting and re-plating (Fig. 4a), a single GFP⁺ colony was picked and expanded for further characterization of naive and primed cells. Integration sites of the reporter in the GFP⁺ colony was determined by splinkerette PCR as described previously³⁰ with slight modification. Genomic DNA (gDNA) was isolated from GFP⁺ hESC_H9, and 1 µg gDNA was digested with DpnII and BfuI overnight, respectively. The digested gDNA was purified with the QIAquick PCR Purification kit (Qiagen), and then ligated to MboI splinkerette linkers overnight. Five microlitres of the ligation reaction product were used for the first round of PCRs with a cycle of 96 °C for 2 min, followed by 10 cycles of 92 °C for 40 s, 60 °C for 40 s and 72 °C for 2 min with a decrease of 1 °C per cycle; 10 cycles of 92 °C for 40 s, 63 °C for 40 s and 72 °C for 1 min with a decrease of 0.5 °C per cycle; 25 cycles of 92 °C for 40 s, 50 °C for 40 s and 72 °C for 1 min; The final elongation was performed for 10 min at 72 °C, and then cooling to 4 °C. The second round of PCR (nested PCR) was done with primers Nested and T-Bal with a cycle of 2 min at 96 °C followed by 6 cycles of 92 °C for 40 s, 66 °C for 40 s and 72 °C for 1 min with a decrease of 1 °C per cycle and 14 cycles of 92 °C for 40 s, 59 °C for 40 s and 72 °C for 1 min. The final elongation was performed for 10 min at 72 °C. Finally, the purified PCR products from the nested PCR were sequenced, showing the same single PCR product under different enzyme digestion. The linkers and primers used in splinkerette PCR are shown in Supplementary Table 1.

Knockout of LBP9 in hESCs. The published CRISPR/Cas9 vector X330³⁸ was modified for the knockout (KO) of LBP9 in this study. Two guide-RNA (gRNA) sequences targeting the second exon of LBP9 were designed according to the guide RNA design tool (<http://crispr.mit.edu/>). gRNA sequences were then synthesized and ligated into the vector of X330 to generate two LBP9 knockout vectors, referred to as CRISPR/Cas9-gRNA(LBP9)#1 and #2. A total of 2.5×10^5 hESC_H9 were transfected with 2.5 µg CRISPR/Cas9-gRNA and 1 µg pT2-GFP, and then seeded onto matrigel/feeder-coated 6-well plates. The cells transfected with Cas9 and pT2-GFP were used as controls. The transfected hESCs were cultured in conventional hESC medium. To enrich for targeted events, GFP⁺ cells were sorted by FACS and re-plated onto matrigel/feeder-coated 6-well plates on day 2 post-transfection. On day 6 post-transfection, single cell suspensions were immunostained with TRA-1-81, and sorted to collect GFP⁺/TRA-1-81⁺ (undifferentiated) and GFP⁺/TRA-1-81⁻ (differentiated) cells, respectively. Genomic PCR was performed on genomic DNA isolated from these undifferentiated and differentiated cells, respectively. PCR products were subjected to TA cloning and sequencing. The gRNA and primer sequences are in Supplementary Table 1.

Gene expression microarrays. Total RNA was isolated from hESCs using the RNeasy kit (Qiagen). The quality of total RNA was checked by gel analysis using the total RNA Nano chip assay on an Agilent 2100 Bioanalyzer (Agilent Technologies). Only samples with RNA index values greater than 8.5 were selected for expression profiling. 100 ng of total RNA was simultaneously processed from each sample. Biotin-labelled cRNA samples for hybridization on Illumina Human Sentrix-12 BeadChip arrays (Illumina, Inc.) were prepared according to Illumina's recommended sample labelling procedure. Data extraction was done for all beads individually, and

outliers were removed when >2.5 MAD (median absolute deviation). All remaining data points were used for the calculation of the mean average signal for a given probe, and standard deviation for each probe was calculated.

RNA-seq. Total RNA was extracted from three types of cells—hiPSCs, HFF-1, EBs differentiated from hiPSCs using Trizol (Invitrogen)—following the manufacturer's instructions. After extraction a DNase treatment was applied using TURBO DNA-free kit (Ambion) and a second RNA extraction with Trizol was performed, and further PolyA(+) RNA extraction and RNA-seq library construction followed Illumina TruSeq RNA Sample Preparation Kit protocol. Sequencing was performed on the Illumina HiSeq 2000 machine with single-end 101 cycles.

Statistical analysis. All data were collected from at least two biological replicates and from at least two independent experiments. No statistical method was used to predetermine sample size. Sample sizes were based on previously published experiments which are similar to the present study. Experiments were not randomized. The investigators were not blinded to the group allocation during the experiments or outcome assessment. All data are shown as mean and standard deviation (s.d.) of multiple replicates/experiments (as indication in figure legends). Analysis of all experimental data was done with GraphPad Prism 5 (San Diego, CA). *P* values were calculated with two-sided, unpaired *t*-test following the tests for differences in variances as specified in the figure legends. *P* values less than 0.05 were considered significant.

Sequencing and mapping. In the pilot study, RNA-seq reads were first filtered by Illumina quality control and then mapped to the human genome (hg19: <http://genome.ucsc.edu/>) by TopHat-1.3.0³⁹ (parameter settings: `-solexa1.3-quals -g 100 -p 4-segment-mismatches 3 -segment-length 30`). Only the aligned reads with unique location in the genome were used for further analysis. At the extended study, we collected 269 samples from 14 independent published studies for pluripotent stem cells (hiPSC and hESC), somatic tissues, cancer cell lines and cells from early embryos (Supplementary Tables 4 and 5). The RNA-seq reads from these published samples and our pilot study were mapped by STAR mapper⁴⁰ (parameter settings: `-readFilesCommand zcat -runThreadN 10 -genomeLoad LoadAndRemove -outFilterMatchNminOverLread 0.66 -outFilterMismatchNoverLmax 0.05 -outFilterMultimapNmax 100`). To control the quality of the data, we only chose the ones with more than half of the total reads being uniquely mapped and the number of uniquely mapped reads larger than 10 million. For mapping details see Supplementary Table 6. For part of the ChIP-seq analysis, the raw sequencing reads were mapped by bowtie2 with default parameter settings⁴¹ and MACS software⁴² was further applied for the peak calling.

Gene expression calculation. Gencode V14 human gene annotation was downloaded from GENCODE Project (<http://www.gencodegenes.org/>). The number of uniquely mapped reads was calculated on each annotated gene, and further normalized to reads per kilobases per million (RPKM) by total number of uniquely mapped reads. At the extended study, featureCounts⁴³ was used for counting the number of uniquely mapped reads at exonic regions of annotated genes.

Expression calculation of repeated elements. The human RepeatMasker annotation file was downloaded from UCSC Tables (<http://genome.ucsc.edu/cgi-bin/hgTables?command=start>), and used as repeat annotation standard in our analyses. The number of reads, uniquely mapped to repeated elements annotated by RepeatMasker, was calculated by featureCounts⁴³, which was further RPKM normalized by total number of uniquely mapped reads. Using uniquely mapped reads, we first calculated the total number of the reads deriving from all repeated elements and each repeat family, respectively. Next we computed the relative abundance and enrichment level of each repeated family. Specifically, the relative abundance of repeated element family A is the percentage of reads allocated to family A, divided by total reads of repeated elements. The enrichment level was calculated using the formula $(N_i \times L) / (N \times L_i)$, where N_i is the number of reads allocated to a specific repeated family, N is the total number of reads allocated to all repeated elements, L_i is the total length of the specific repeated family and L is the total length of all repeated elements. To determine the relative abundance and enrichment of LTR elements, we applied the above strategy, except reads of all LTR elements were used instead of all repeated elements. One-tail binomial test was applied as a statistical tool.

To determine the expression level of HERVH, full-length HERVH was defined as LTR7-HERVH-int-LTR7. First, RepeatMasker was used to annotate all repeated elements, and HERVH-int and LTR7 terminals were mapped to the whole human genome (hg19). Then, the distribution of the distances between HERVH-int and neighbour LTR terminal fragments was calculated, and the HERVH-int and LTR terminal elements within the 99% quantile of the distance distribution (2,655 bp) was further merged. The median size of the full-length HERVHs was found to be 5,750 bp. Using the above strategy, 1,225 full-length HERVHs were identified in total, including 1,057 elements with LTRs at both ends (DiLTR), 159 HERVHs with one terminal LTR (monoLTR) and 9 HERVHs with no recognizable LTR (NoLTR) (Supplementary Table 7). The expression and enrichment level of full-length HERVHs was calculated by the same procedure as above. To define the transcriptionally active

and inactive loci of HERVHs in hPSC samples, we analysed 1,225 full-length HERVH elements by the hierarchical cluster analysis. The hierarchical distances among samples were based on Spearman's correlation coefficient. To minimize the total within-cluster variance the hierarchical distances among full-length HERVHs were calculated by the Euclidean distance with Ward's method. All calculation was based on raw normalized expression value (RPKM). To visualize the expressed HERVH elements, HERVHs with expression levels with or above 8 RPKM were capped to 8, while the ones equal to or below 0.125 were treated as 0.125. During logarithmic transformation process a small number (0.01 RPKM) was added to the expression level of all the genes or repeated elements to handle instances of zero expression.

Identification and characterization of HERVH-derived chimaeric transcripts and HERVH neighbouring genes. The search for HERVH-derived chimaeric transcripts in hPSCs was done by looking for the junction reads that have one part mapped to the exon-free full-length HERVH region and another part mapped to the exonic region of annotated protein-coding genes. The expression level of chimaeric transcripts was quantified by counting the number of reads sharing the same chimaeric junction. Chimaeric transcripts supported by at least 10 junction reads were used for analysing samples from inter-cell-type comparison (Supplementary Tables 8 and 9). The neighbouring gene of HERVH is defined as the closest gene(s), while HERVH-derived genes are the ones whose exonic regions overlap with HERVH. To determine the transcription start site (TSS), we re-analysed the published hESC_H1 CAGE data from the ENCODE project. The relative location of TSSs on active HERVH elements was profiled. We calculated (1) the density distribution of CAGE fragments around HERVHs, and (2) their relative position in LTR7-HERVH-int-LTR7. The positive value of the peak indicates that TSS is mainly located at the HERVH-LTR boundary regions (Extended Data Fig. 4c).

ChIP-seq comparative analysis. Global hESC_H1 chromatin statuses based on HMM method was proposed by Ernst *et al.*⁴⁴ and was downloaded from ENCODE (<https://genome.ucsc.edu/ENCODE/>). Then, ChIP-seq peak files and bigWig files for H1 DNaseI hypersensitivity and histone modification information were also downloaded from the same source. Furthermore, bigWig files for H3K9me3, H3K27me3 and H3K4me3 in penis foreskin fibroblast primary cells, H1-hESC and hiPSCs were downloaded from Epigenome Atlas (<http://www.genboree.org>) for inter-cell type comparison. In the comparison of histone modification between naive-like stem cells and primary stem cells, the peak files provided by Gafni *et al.*⁴ and the raw sequencing data provided by Chan *et al.*³ were downloaded from the corresponding sources, and their processing is described in the sequencing and mapping sections. Bwtools (<https://github.com/CRG-Barcelona/bwtool/wiki>)⁴⁵ was applied for facilitating bigWig file processing, where aggregate function was used for the calculation of average ChIP-seq signal surrounding given regions and matrix function was used for ChIP-seq signal detection around each given region. In the comparative study of ChIP-seq peak enrichment analysis (Fig. 2a and Extended Data Figs 2a and 9f), the ChIP-seq peaks within 10 kb of HERVH centres were kept for the analysis, and the distances of these peaks to the closest HERVH boundaries were calculated, where the mean difference between the distances for active ones and inactive ones was compared by Student's *t*-test. At the same time, the number of active HERVHs or inactive ones containing ChIP-seq peaks within 10 kb of their centres was calculated, and two-sided binomial test was applied for the significance calculation of peak enrichment in active ones. In the comparative study of the difference of ChIP-seq coverage distributions between active HERVHs and inactive ones, the areas within 10 kb of HERVH boundary were considered, and the coverage levels for different loci within this region were calculated in continuous 10-kb windows.

Transcription factor analysis. To identify candidate transcription factors binding HERVH we took *in silico* and data mining approaches. *In silico*: CLOVER⁴⁶ was used to compare active HERVHs against GC matched control using the JASPAR core vertebrate motifs (http://jaspar.genereg.net/cgi-bin/jaspar_db.pl?rm=browse&db=core&tax_group=vertebrates). GC matched controls were 20-kb sections of the human genome 5' of known genes and within 0.05% of the GC content of the focal sequences. Using ROVER⁴⁷ we determined motifs enriched in the more active HERVHs, those with LTR7, compared with those that are active but less so (those with LTR7C/Y). In addition we compared the standard version of LTR7 (seen in HERVH) against the less active HERVH sequences and compared the active HERVH sequences with HERVK active sequences (Extended Data Fig. 3b). OCT4 and NANOG ChIP-seq data³ in hESC_H1 were downloaded from ArrayExpress (E-MTAB-2044). The raw sequencing reads were mapped to human genome (hg19) by bowtie2 with default parameter settings⁴¹, and MACS software⁴² was further applied for the peak calling.

DHS analysis. ENCODE project⁴⁸ DHS files were downloaded in bed format. The 'closest' method in Bedtools⁴⁹ was used to find overlapping or the closest DHSs. To investigate the statistical significance of the number of sequences including one or more DHSs, we conducted a Monte Carlo simulation. According to the transcriptionally active HERVHs, we generated random sequences of the same length on the same chromosome and then counted the number of sequences including DHSs.

We repeated this 10,000 times and counted how many of iterations included more or the same number of DHSs than observed in our active HERVH sequences (none). To enable accurate estimation of type I error rate, we define $P = (n + 1)/(m + 1)$, where n is the number of observations as or more extreme than observed and m the number of trial runs. A vicinity of 1.5 kb on both sides of sequences was also searched for DHS. We used chi-square to compare observed number of inactive sequences overlapping one or more DHS with the number we would expect if there was no difference between the two.

Analysis of chromatin marks and DNA methylation. The methylation profiles of H3K4me3 and H3K27me3 in hESC_H7 are available at the ENCODE portal. We focused on the data sets generated by standard protocols. We compared averages for histone marks H3K4me3 and H3K27me3 on active and inactive HERVHs and also LTR7. We counted the number of methylation sites reported for each group and kept the extension size 1.5 kb consistent with DNase analysis.

We also compared CHD1's binding sites in active and inactive extended HERVH. CHD1 binding sites in ESCs were downloaded from ENCODE (<http://genome.ucsc.edu/cgi-bin/hgFileUi?db=hg19&g=wgEncodeSydhTfbs>, accessed on 7 December 2012). HERVH sequences were extended 1,500 bp on both sides and the number of CHD1's binding sites overlapping the extended sequences determined. Chi-square test was employed to test for significance. A similar method as the one explained for histone methylation analysis was used to calculate the expected value. We also compared binding sites of MYC, MAX and CHD2 chromatin remodellers, available through the ENCODE portal (<http://genome.ucsc.edu/cgi-bin/hgFileUi?db=hg19&g=wgEncodeSydhTfbs>, Release 3, accessed on 7 December 2012). Using the same approach as above we compared active and inactive extended HERVH, its LTR7 and also HERVK and its LTR5.

To study the global DNA methylation status of HERVHs in hPSCs, we downloaded the genome-wide bisulphite sequencing data in wig format from Epigenome Atlas (<http://www.genboree.org/epigenomeatlas/index.rhtml>) for hiPSCs, H1s and penis foreskin fibroblast primary cells (see Supplementary Table 4). We used BEDtools⁴⁹ (<https://code.google.com/p/bedtools/>) to extract the methylation scores for detected CpGs in each HERVH-associated LTR7, and then calculated the average methylation level for each LTR7. To compare DNA methylation status differences of HERVH-associated LTR7s in hPSCs versus fibroblast cells, we applied one-sided Wilcoxon rank sum test.

Estimating the coding potential of the HERVH-driven ncRNAs. We established a set of putative ncRNAs that appear to be HERVH associated. For each of these we queried LNCipedia⁵⁰ (<http://www.lncipedia.org/>) via gene name, or if that failed, via transcript ID. If present this resource reports Coding Potential Calculator (CPC) scores⁵¹, possible Pfam motifs and presence in the PRIDE database (a database of mass spectrometry identified proteins including small peptides). As all of the sequences are PRIDE negative we don't report this. In the few instances where the transcript was unknown to LNCipedia we determined CPC and Pfam scores via the CPC website (<http://cpc.cbi.pku.edu.cn/>). CPC values under zero are considered evidence for non-coding potential. Scores between 0 and 1 are weak candidates for coding function. Scores over one are considered as stronger evidence for coding. Nine of the RNAs have negative CPC scores (meaning most likely to be ncRNA), 18 have scores between 0 and 1 (possibly with small fragment that might be protein coding) and 7 have scores over 1 (meaning that they are more likely to have coding potential) (Supplementary Table 11).

HERVH-derived lncRNAs and shHERVH targeting prediction. We searched HERVH-derived lncRNAs by looking for the lncRNAs with exonic regions overlapping with hPSC-specific full-length HERVHs (Supplementary Table 10). The annotation of lncRNAs was downloaded from Gencode V14 (<http://www.genencodegenes.org/>). Using the sequences of the shHERVH constructs, used in the knockdown experiments (shHERVH#3, shHERVH#4 and shHERVH#12), we predicted their targets (21-bp perfect matching). Next, we identified genes that either form chimeric transcripts with the targeted HERVHs or are derived from them. Using our global gene expression profiling data (Illumina), we also examined if any of these genes are significantly downregulated (one-sided Student's *t*-test, *P* values adjusted by Benjamini and Hochberg method).

Global gene expression analysis. Expression data was processed from bead-level expression intensity values pre-processed from Illumina's software in the form of .txt or .bab files carrying 48,324 probe sets targeted by HumanHT-12 v4 Expression BeadChips. Green intensities were extracted after adjusting non-positive values by BeadArray's (<http://bioconductor.org/biocLite.R>) built in functions. Furthermore, to the BeadArray output data, we fetched significance level of normalized expression values corresponding to probe ID using the lumi (<http://www.bioconductor.org>) package. We exploited inbuilt functions of this package such as variance-stabilizing transformation (VST) to deal with sample replicates and robust spline normalization (RSN), for normalization, of which *P* value <0.05 were further transformed onto log₂ scale. Probe IDs were annotated by using Bioconductor package (<http://www.bioconductor.org/packages/release/data/annotation/html/illuminaHumanv4>.

db.html). Expression values of multiple probes for one gene were assigned by their median, resulting in 20,394 unique genes for GFP-marked samples.

In this study, fold-change of differential expression between samples on log₂ scale was analysed using linear and Bayesian model algorithms from limma (<http://bioconductor.org/>) and pairwise differential expression between samples from various data sets was performed by the correction of batch effect arising from two different platforms by normalizing (quantile) each data set to a sample of the same genotype and merging data sets for downstream analysis. Heat maps (Fig. 3e) shown for differential expression among LBP9 knockdown and HERVH knockdown (shLBP9 and shHERVH) and control (shGFP) samples were drawn for genes, showing significantly highest standard deviations, on their Z-score. The matrix was hierarchically clustered (Spearman correlation and distances between observations were calculated using euclidean distances and average linkage). We exploited the online tool GOrilla (<http://cbl-gorilla.cs.technion.ac.il/>) to check for biological processes functional enrichment (Extended Data Fig. 9j) of differentially expressed genes where the entire gene list was used as background. A false-discovery-rate-corrected *P*-value threshold was set at 0.05.

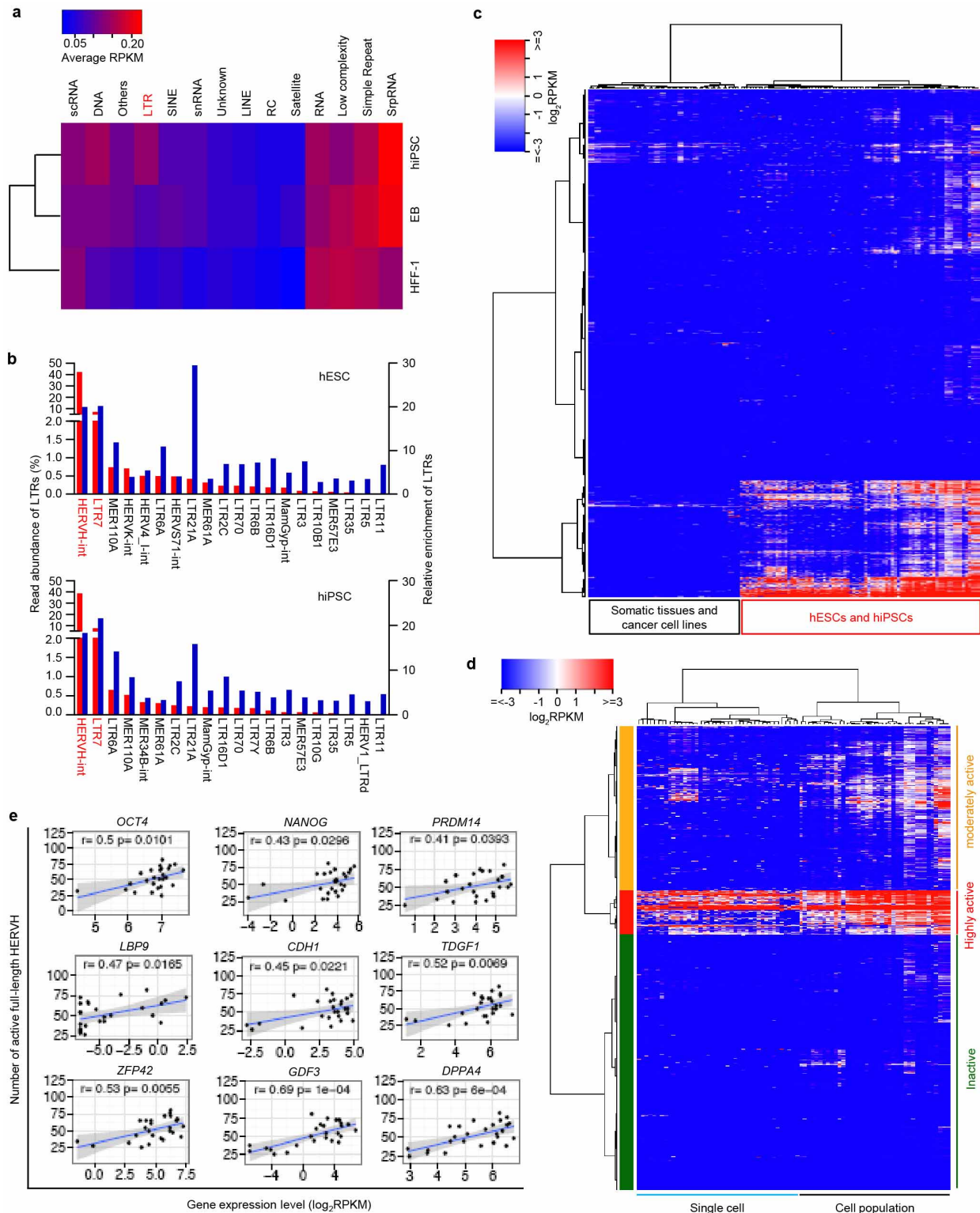
Comparison of global expression profile of human ICM, hESC⁵² (GSE29397) and GFP-marked samples (present study) represented gene-wise (19,103 genes possessing common probes between two platforms) were subjected to hierarchical clustering (Pearson correlation, centroid linkage, $k = 3$), whereas samples were clustered using Spearman correlation, centroid linkage and height represent the units of euclidean distance. In Fig. 4e, g, units for height on y-axis represent the distance (*D*), which is the value of criterion associated with method by which they are clustered ($D = 1 - C$; *C*, correlation between spot clusters). Differentially expressed gene-list between GFP^{high} and GFP^{low} samples (FDR <0.05) were intersected to cross-platform, pairwise comparison of rescaled expression values of genes assigned as their row-wise Z-score (expression value subtracted by mean of its row values and divided by its standard deviation). Neighbouring genes were fetched using bedtools falling in the window of 50 kb from HERVH genomic coordinates, fold changes between naive and primed were calculated independently, keeping thresholds for human and mouse samples in the same way as mentioned above, data sets were intersected by gene names, and heat maps were drawn on their calculated Z-scores.

Cross-species gene expression analysis (see ref. 4) was performed on human with Illumina HumanHT-12 v4 (expression beadchip containing 47,324 probes, present study) and Affymetrix HuGene 1.0 ST microarrays (containing 33,252 probes, GSE46872) and on mouse with Agilent 4x44K array platform (containing 45,018 probes, GSE15603) microarray expression sets. Human-mouse orthologous genes were downloaded with an online tool (biomart) from Ensemble (<http://www.ensembl.org/biomart/martview/>) containing 18,657 pairs of orthologous genes; out of these a total of 9,583 genes were mapped by probes of both human and mouse array platforms and explored in the present study, which were implemented for further analysis. The expression value of each gene was determined by median of all probes targeting to it. As mentioned above, the batch effect was corrected; correction was confirmed by Principal Component Analysis (PCA). Next, these independent data sets were merged in one for further analysis. Each gene value was further assigned as their relative abundance value, which is the expression value of a gene in each sample divided by the mean of expression values of corresponding genes across the samples within the same species. The resulting expression matrix (Fig. 4f) was subjected to hierarchical clustering (Spearman's correlation, average linkage); the *P* value threshold for the correlation test for the matrix was kept up to 0.01. Whereas outliers are not shown in the coloured matrix, the hierarchically clustered dendrogram displays all of the samples included in the analysis.

Comparative analysis of primed and naive-like hESCs to human ICM. To compare GFP^{high}, GFP⁺ and GFP^{low} hESCs with human ICM, human ICM data⁵² were re-analysed along with previously described naive and primed samples^{4,53}. These data sets were generated on different platforms, so they were subjected to the same pre-processing. In brief, we fetched 19,102 common genes probed on all the platforms, with the value of individual gene denoting the mean of its expression value. The batch effect resulting from two different platforms was removed by quantile normalization of each data set to a sample of the same genotype which was then excluded from analysis. Additionally, batch effect arising from ICM data was corrected by quantile normalization to the mean values of its ESC samples, which enabled it to be consistent with the normalized data sets of GFP, naive and primed samples. The samples were hierarchically clustered using average linkage and Spearman correlation as a distance matrix via multi-scale bootstrap resampling, replicated 1,000 times. Moreover, *P* values were computed for each of the clusters by approximately unbiased and bootstrap probability, which enabled us to assess the uncertainty in hierarchical cluster analysis. Outlier samples (approximately unbiased and bootstrap probability <50%) are not shown in the plot (Fig. 4e) but were included throughout statistical analysis.

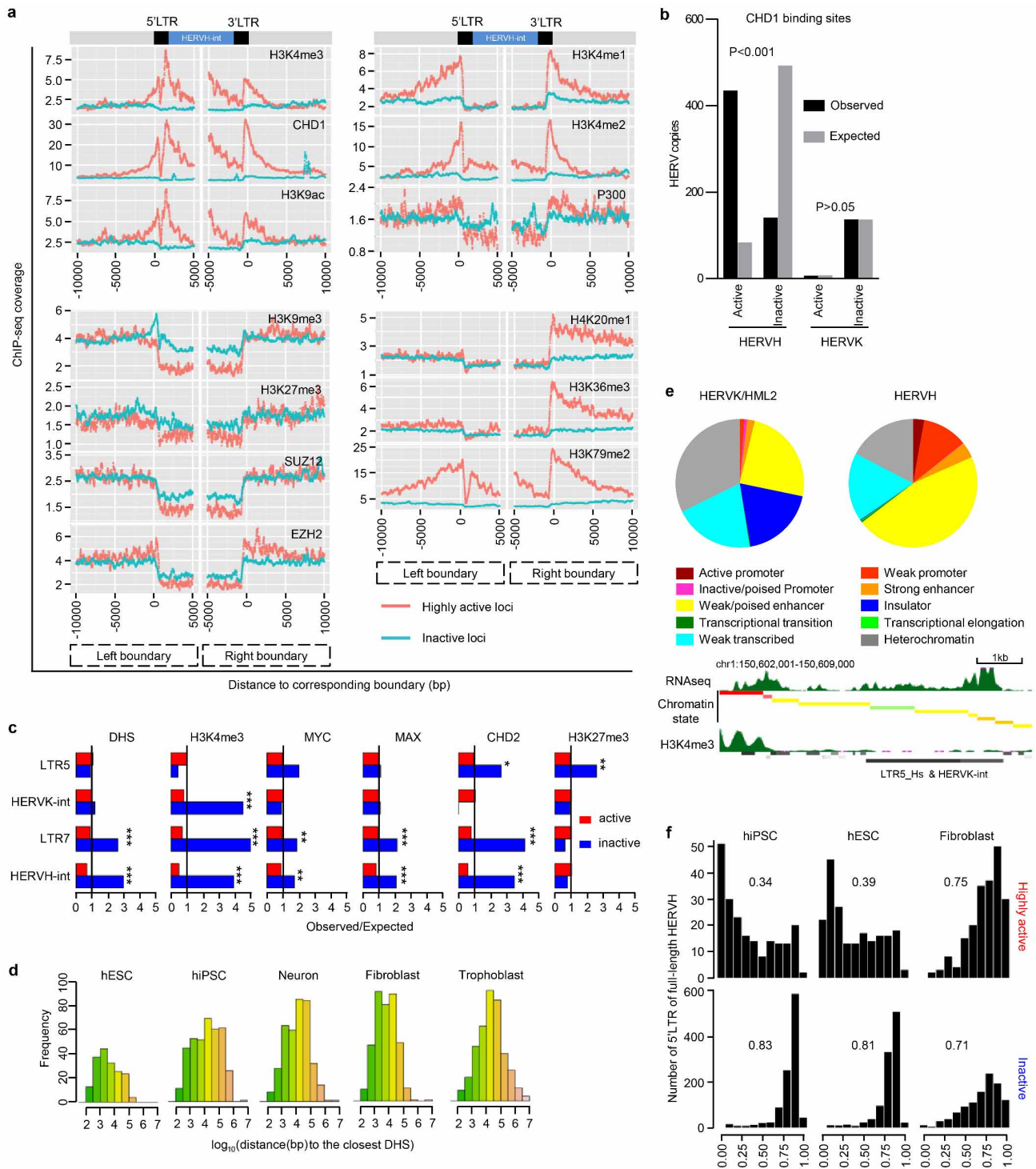
Graphics. Graphics (for bioinformatics analyses) were done in R (<http://www.r-project.org/>), and ggplot2 (<http://ggplot2.org/>) was used for part of the graph making.

30. Grabundzija, I. *et al.* Sleeping Beauty transposon-based system for cellular reprogramming and targeted gene insertion in induced pluripotent stem cells. *Nucleic Acids Res.* **41**, 1829–1847 (2013).
31. Haase, A. *et al.* Generation of induced pluripotent stem cells from human cord blood. *Cell Stem Cell* **5**, 434–441 (2009).
32. Prigione, A., Fauler, B., Lurz, R., Lehrach, H. & Adjaye, J. The senescence-related mitochondrial/oxidative stress pathway is repressed in human induced pluripotent stem cells. *Stem Cells* **28**, 721–733 (2010).
33. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
34. Onder, T. T. *et al.* Chromatin-modifying enzymes as modulators of reprogramming. *Nature* **483**, 598–602 (2012).
35. Ivics, Z., Hackett, P. B., Plasterk, R. H. & Izsvak, Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **91**, 501–510 (1997).
36. Kaufman, C. D., Izsvak, Z., Katzer, A. & Ivics, Z. Frog Prince transposon-based RNAi vectors mediate efficient gene knockdown in human cells. *J. RNAi Gene Silencing* **1**, 97–104 (2005).
37. Wang, Z., Oron, E., Nelson, B., Razis, S. & Ivanova, N. Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell* **10**, 440–454 (2012).
38. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
39. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
40. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
41. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
42. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
43. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
44. Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nature Biotechnol.* **28**, 817–825 (2010).
45. Pohl, A. & Beato, M. bwtool: a tool for bigWig files. *Bioinformatics* **30**, 1618–1619 (2014).
46. Frith, M. C. *et al.* Detection of functional DNA motifs via statistical over-representation. *Nucleic Acids Res.* **32**, 1372–1381 (2004).
47. Haverly, P. M., Hansen, U. & Weng, Z. Computational inference of transcriptional regulatory networks from expression profiling and transcription factor binding site identification. *Nucleic Acids Res.* **32**, 179–188 (2004).
48. Nepf, S. *et al.* An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* **489**, 83–90 (2012).
49. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
50. Volders, P.-J. *et al.* LNCipedia: a database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Res.* **41**, 246–251 (2012).
51. Kong, L. *et al.* CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* **35**, W345–W349 (2007).
52. Vassena, R. *et al.* Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development. *Development* **138**, 3699–3709 (2011).
53. Hanna, J. *et al.* Metastable pluripotent states in NOD-mouse-derived ESCs. *Cell Stem Cell* **4**, 513–524 (2009).
54. Bellucci, M., Agostini, F., Masin, M. & Tartaglia, G. G. Predicting protein associations with long noncoding RNAs. *Nature Methods* **8**, 444–445 (2011).
55. Zhou, W. *et al.* Induction of human fetal globin gene expression by a novel erythroid factor, NF-E4. *Mol. Cell. Biol.* **20**, 7662–7672 (2000).
56. Havugimana, P. C. *et al.* A census of human soluble protein complexes. *Cell* **150**, 1068–1081 (2012).



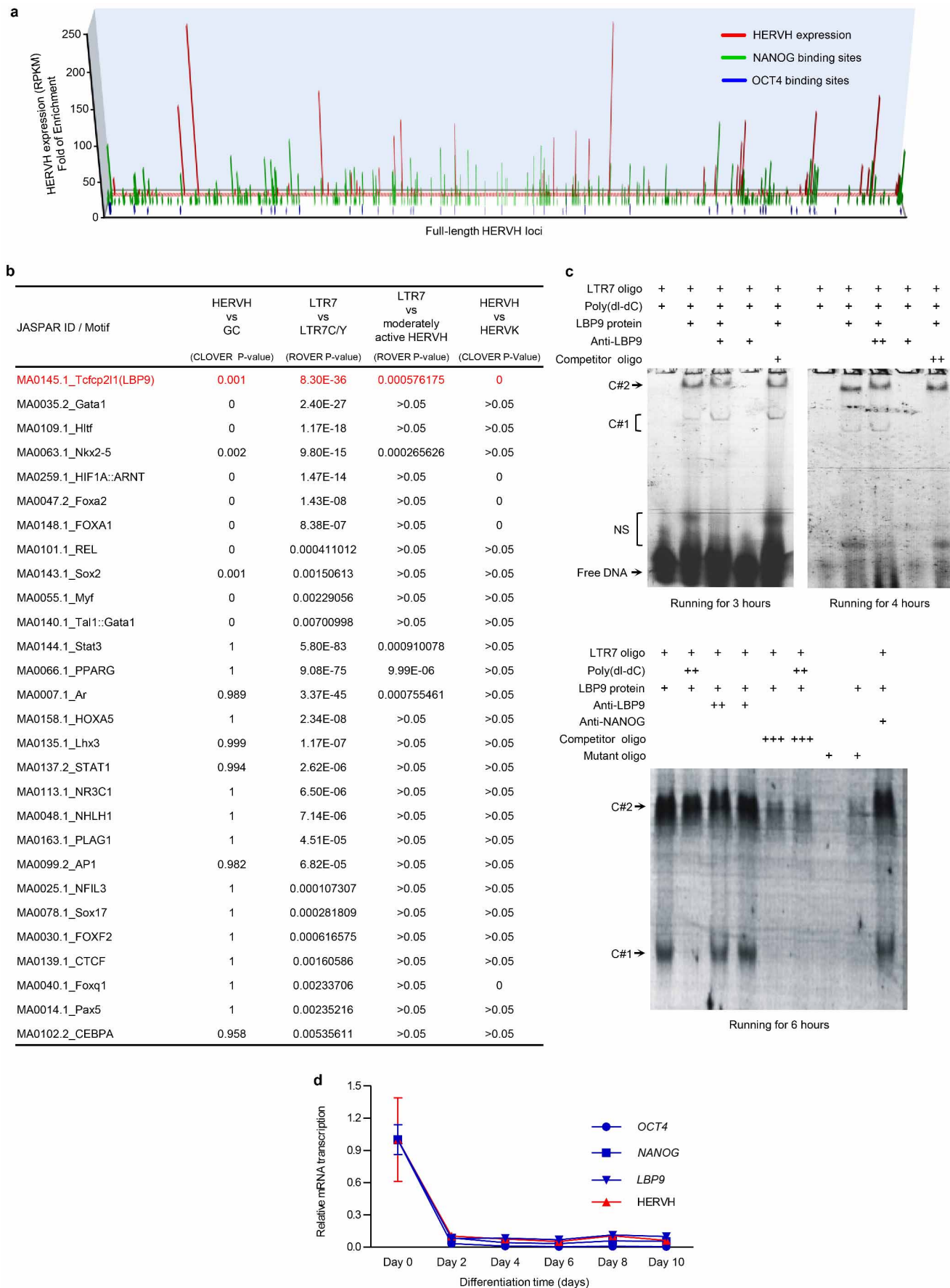
Extended Data Figure 1 | HERVH is the most transcriptionally enriched transposable element in hPSCs. **a**, Heat map showing expression of repetitive element classes in human induced pluripotent stem cells (hiPSCs), fibroblasts (HFF-1) and hiPSC-derived embryoid bodies (EBs). **b**, Highly expressed (top 20) LTR elements in hESCs (top panel) and hiPSCs (bottom panel). Red bars: proportion of reads of each LTR element in total LTR-element-related reads. Blue bars: enrichment of each LTR element relative to the background (calculation details described in Methods). **c**, **d**, Heat maps showing the expression profile of 1,225 full-length HERVHs in various human cell types. For a list of samples and expression data see Supplementary Tables 4 and 7, respectively. **e**, Expression profile of HERVH in 43 normal somatic tissues,

8 cancer cell lines and 55 hESC (H1, H6 and H9) and 26 hiPSC samples, including our hiPSC³⁰ line. The rows represent the transcription from 1,225 full-length HERVH loci. **d**, Expression profile of HERVHs in hPSC lines and single cells from three individual hESC clones. On the basis of their expression, the 1,225 full-length HERVH loci are clustered into three groups (highly, moderately and inactive). Note that HERVH activity is heterogeneous between single cells of an hPSC population. **e**, HERVH expression in single hESCs positively correlates with the expression of key pluripotency-associated transcription factors. Note that SOX2, not illustrated, shows no correlation ($P = 0.59$). Each dot represents a single hESC sample²⁴.



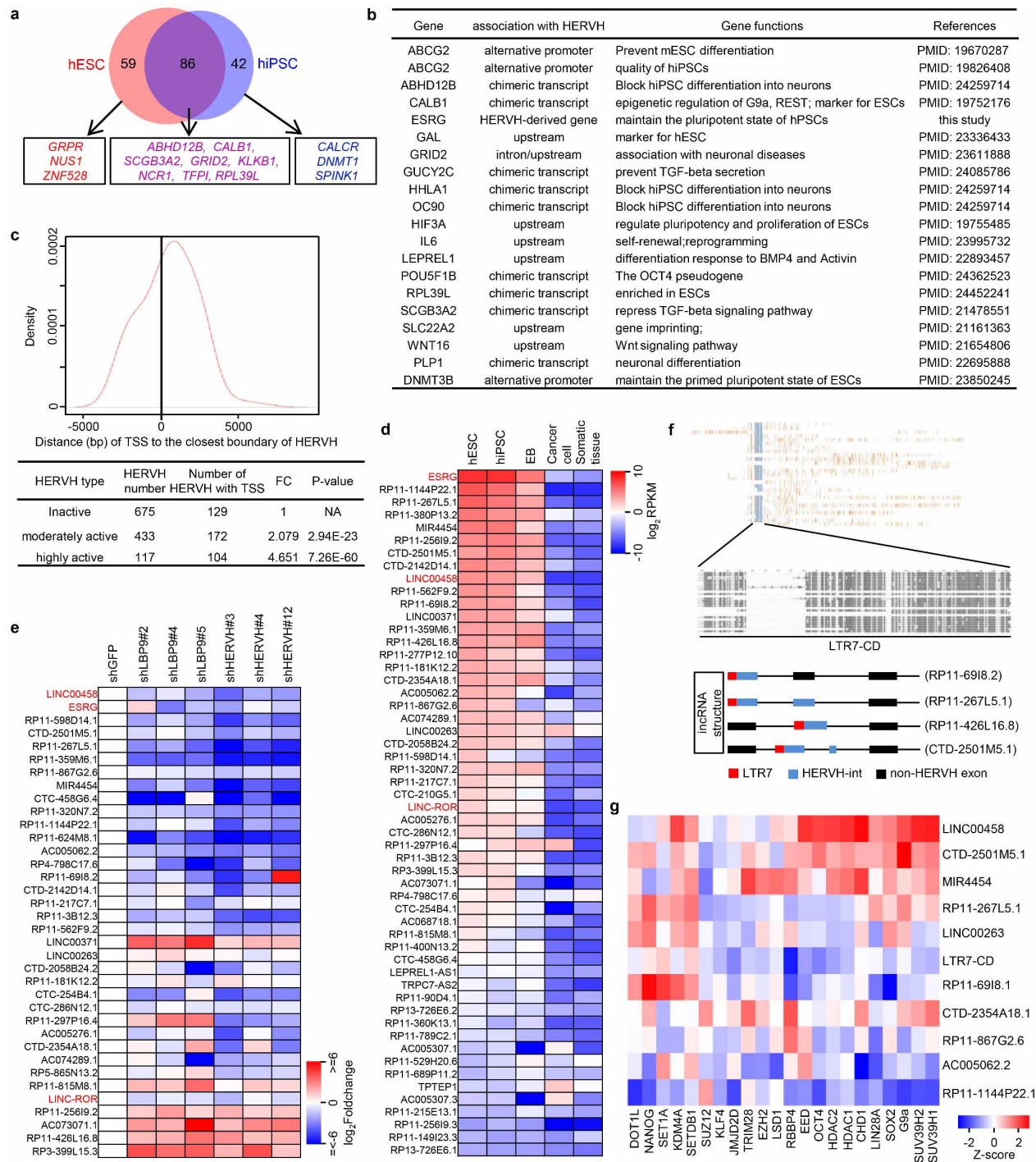
Extended Data Figure 2 | HERVH shows the hallmarks of active chromatin in hPSCs. **a**, Chromatin status analysis around full-length HERVHs in hESC_H1. The promoter/transcription initiation regions and the transcribed regions of active HERVH loci are associated with active epigenetic marks and chromatin modifiers. The neighbouring regions of inactive HERVH loci show the hallmarks of heterochromatin. **b**, Active HERVHs were enriched with binding sites for CHD1 compared to inactive ones. Chi-squared tests were performed, P values shown as statistical significance. **c**, Comparison of epigenetic marks and chromatin modifiers in proximity of HERVH internal sequence (HERVH-int) and LTR7. As a control, we employ HERVK-int and LTR5. We compare the number of marks within or near active and inactive versions (allowing 1.5 kb either side) of each element in ES cells. Expected numbers are derived from a null of no relative enrichment and P values determined by Chi-squared. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (for data see

Supplementary Table 15). **d**, Cross-tissue comparison of the distance of the closest DHS to the active sequences not including any DHS. The distances are presented in log ratio. **e**, Pie charts show chromatin state segmentation for hESC_H1 in full-length HERVK/HML2 and HERVH regions. Most HERVK regions are repressed while a sub-population of HERVH loci is active. Chromatin status analysis of HERVK/HML2 loci reveals that transcription of the few activated HERVK loci is promoted primarily by neighbouring regulatory elements, and not by their own LTRs. The chromatin status of a representative locus is shown (bottom panel). **f**, Whole-genome bisulphite sequencing analysis on LTR7s. Comparison of the DNA methylation status of actively transcribing (highly active) and inactive elements in three different cell types: hiPSCs, hESCs and fibroblast. Average methylation levels are shown. Data from the ENCODE project and Epigenome Atlas (Supplementary Table 4).



Extended Data Figure 3 | Pluripotency-associated transcription factors bind to HERVH. **a**, All 5' LTR7s of active HERVHs are associated with NANOG, while OCT4 is present in around 39. The plot combines the expression values of the 1,225 full-length HERVH (RNA-seq) with the fold-enrichment values of ChIP-seq data of OCT4 and NANOG in hESC_H1 (ref. 3). Each data point reflects a single full-length HERVH element. **b**, Motifs found to be significant in CLOVER and ROVER analyses. The four comparisons are active HERVH versus GC matched control sequence, HERVH flanked by LTR7 versus those flanked by LTR7C/Y, LTR7 itself against less active HERVH, and active HERVH versus active HERVK. We include only instances where the first two analyses both reported significance. Results for Tfcpl1 (also called LBP9) are shown in red. **c**, EMSA confirms the binding of LBP9 to LTR7 sequence *in vitro*. Two different complexes (C#1 and C#2) were detected in the presence of nonspecific competitor (poly(dI-dC)). Complex 1 has lower stability (adding equal amount of competing oligonucleotide to the binding reaction doesn't destroy it, but 100× excess does). Supershift is not detected with adding anti-LBP9 antibody, suggesting

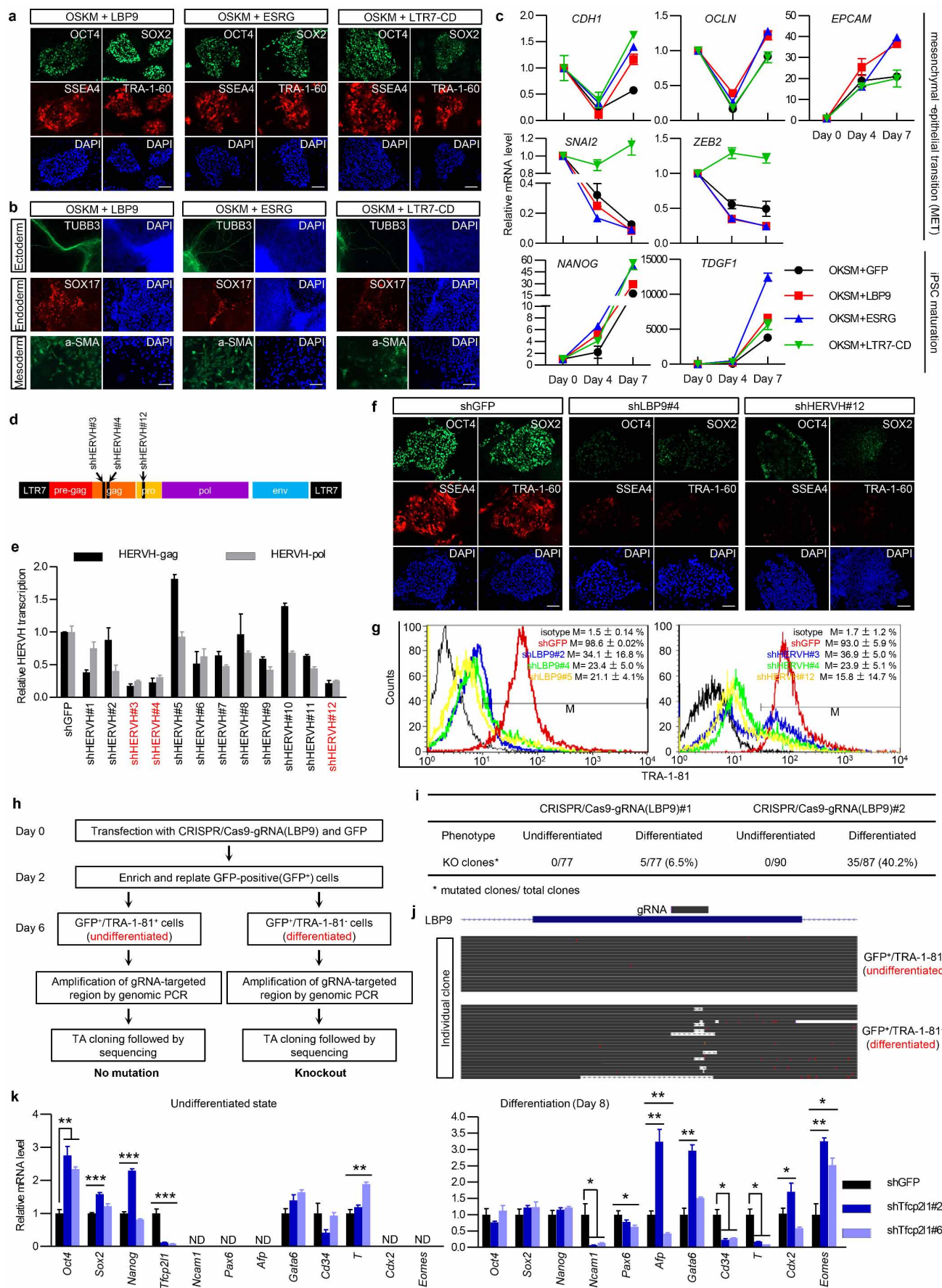
a lack of specificity, at least under our conditions. Complex 2 is resistant to being challenged with the competing oligonucleotide (100-fold excess), and supershifts with anti-LBP9 antibody, indicating specificity. From the low mobility we suspected that complex 2 is a large multimeric complex—this would also account for the modest but reproducible supershift. To explore the potentially multimeric nature of complex 2, we added anti-NANOG antibody. The supershift with anti-NANOG indicates that LBP9 binds LTR7 in a complex with NANOG. *ESRG*-oligonucleotide 50 nM (+); poly(dI-dC) 450 ng (+), 900 ng (++) ; anti-LBP9 5 µg (+), 10 µg (++) ; anti-NANOG 5 µg; competitor oligonucleotide 5 nM (+), 500 nM (++) , 5,000 nM (+++) ; mutant oligonucleotide 50 nM; LBP9 ~10 µg crude extract lysate in 20 µl total reaction volume. NS, nonspecific complex. For a list of oligonucleotide sequences, see Supplementary Table 1. **d**, Relative mRNA expression levels of HERVH correlates with pluripotency-associated transcription factors (*OCT4*, *NANOG*, and *LBP9*) during *in vitro* differentiation of hiPSCs. mRNA levels are normalized to *GAPDH* and relative to day 0. Error bars indicate s.d. from three independent cell cultures per time point.



Extended Data Figure 4 | HERVH-driven transcription in hPSCs.

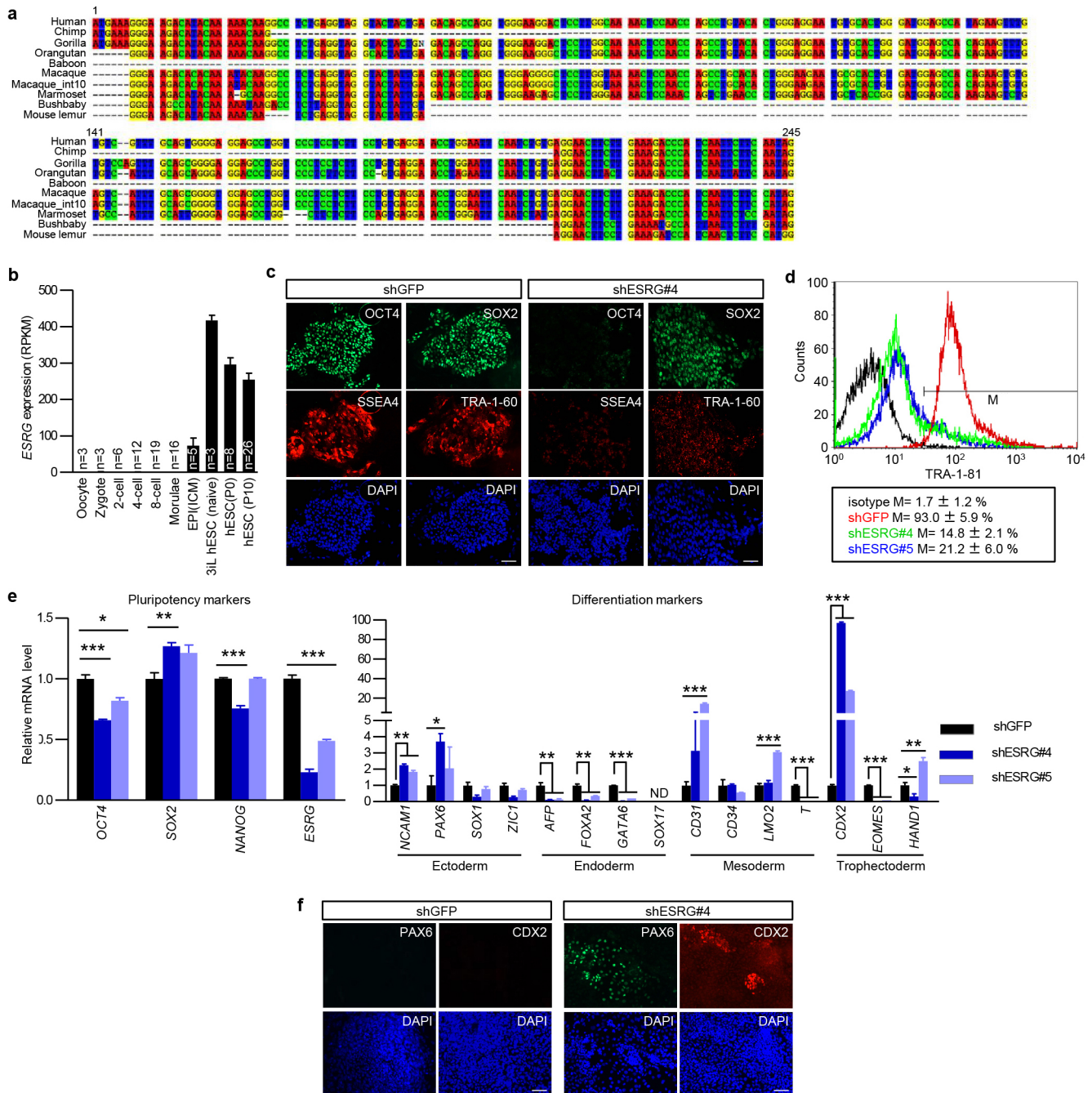
a, HERVH affects the neighbouring gene expression and produces HERVH-specific 'chimeric' transcripts (RNA-seq reads which span HERVH and coding exons of neighbouring genes). Venn diagram shows the overlap between affected genes (see also Supplementary Tables 8 and 9). Examples of genes from each category are shown in boxes. **b**, Genes associated with HERVH function in stem cells with previously described gene functions. **c**, TSS distribution around HERVHs and the relationship between TSS identification and gene activity. CAGE data (from ENCODE) were analysed to identify TSS enriched on 5' end active HERVHs. **d**, Expression heat map of 54 HERVH-derived lncRNAs in hPSCs and differentiated cells. Analysis of RNA-seq data sets as in Extended Data Fig. 1c. Data are displayed as log₂ RPKM with high and low expression shown in red and blue, respectively. EB, embryoid body (data from this study). **e**, Knockdown effects of *LBP9* and HERVH on the highest expressed lncRNAs in hPSCs (selected from the list presented in **d**). mRNA levels are normalized to *GAPDH*, and relative to shGFP expressing,

undifferentiated hESC_H9. Fold-change values relative to shGFP knockdown are shown. Note that the knockdown effects of *LBP9* and HERVH are highly similar. **f**, Alignment of top 22 hPSC-specific/HERVH-derived lncRNAs predict a conserved core domain (CD, referred as LTR7-CD). Certain CDs, embedded within lncRNAs, are annotated as exons and predicted to have limited coding potential (see also Supplementary Table 11). **g**, Heat map of potential RNA-protein interactions (predicted by CatRAPID⁵⁴). lncRNAs were selected from Extended Data Fig. 4f if they were: (1) highly expressed in hESCs; (2) downregulated in HERVH knockdown; (3) enriched in nucleus. The Z-score describes the deviation of the studied RNA-protein interaction propensity from the ones based on randomized 100 RNAs against randomized 100 proteins (calculated by CatRAPID). The core domain of HERVH-derived lncRNAs is predicted to bind RNA-binding proteins, including pluripotency factors (for example, NANOG), and histone modifiers (for example, SET1A and SETDB1). High and low interaction potentials are shown in red and blue, respectively.



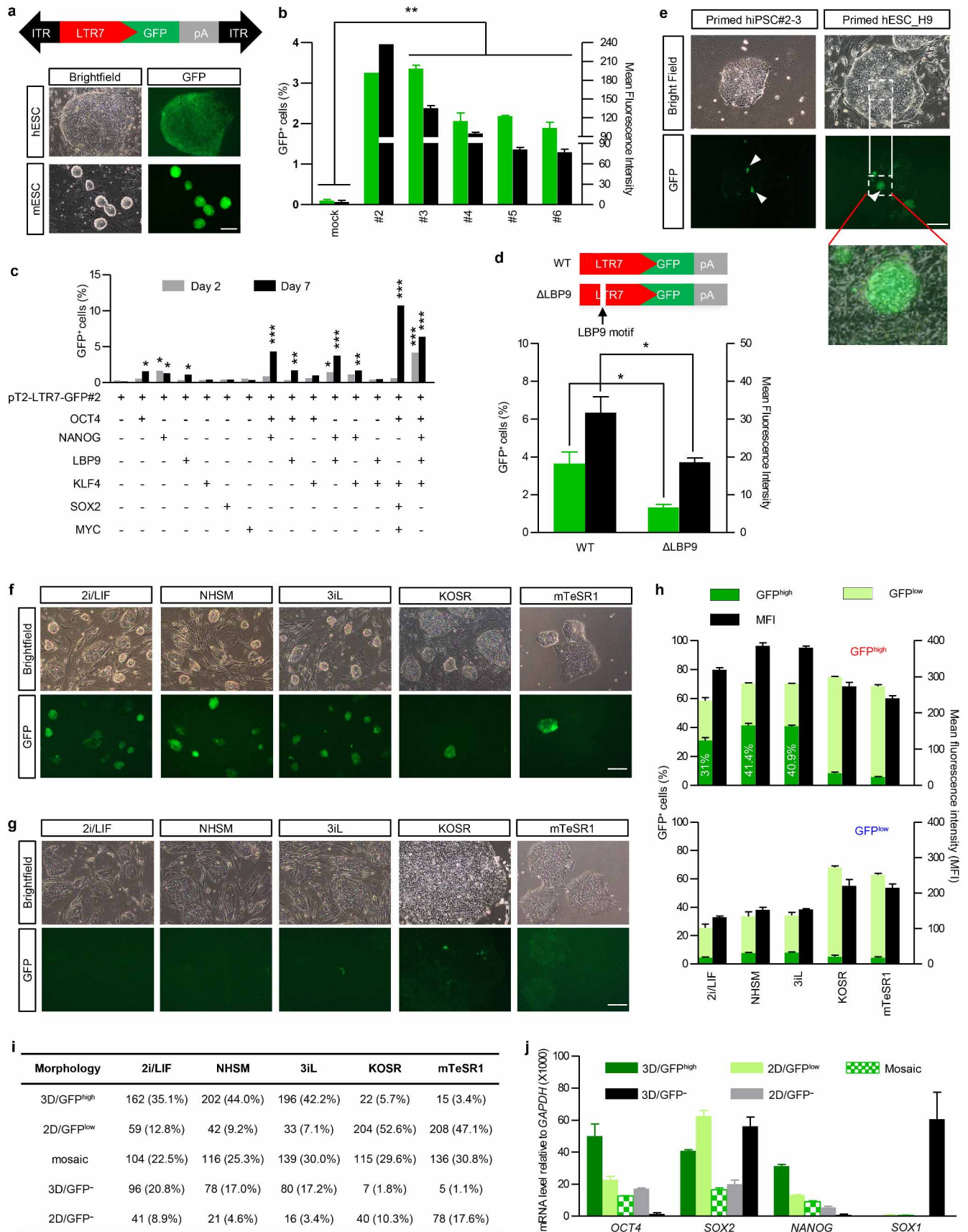
Extended Data Figure 5 | LBP9/HERVH-driven transcription regulates pluripotency in hPSCs. **a, b**, Characterization of hiPSC lines induced by OSKM+*LBP9*, OSKM+*ESRG* and OSKM+LTR7-CD by immunostaining (scale bar, 100 μ m). **a**, Immunostaining for pluripotency markers. **b**, hiPSCs induced by OSKM+*LBP9*, OSKM+*ESRG* and OSKM+LTR7-CD can be differentiated into three germ layer lineages *in vitro*. **c**, Relative expression values of reprogramming-associated genes in HFF-1 are shown at different time points (RT-qPCR). Data normalized to *GAPDH*, and relative to day 0. Error bars indicate s.d. ($n = 3$ independent experiments with biological triplicates per experiment). **d**, Schematic representation of the regions of HERVH targeted by shRNA constructs shHERVH#3, shHERVH#4 and shHERVH#12. Predicted direct targets of shRNAs are shown in Supplementary Table 14. **e**, Validation of the shHERVH constructs. Stable, G418-resistant hESC-derived colonies express various shRNA constructs, targeting HERVH. Knockdown effect was monitored by qRT-PCR detecting either HERVH-gag or HERVH-pol levels. Data shown are representative of two independent experiments with biological triplicates per experiment. shHERVH#3, shHERVH#4 and shHERVH#12 knocked-down $\sim 80\%$ of HERVH compared to the control shGFP. shHERVH#3, shHERVH#4 and shHERVH#12 (all shown in red) are also used in experiments presented in Fig. 3c–f. **f**, Representative immunostaining images showing reduction of pluripotency markers (OCT4, SOX2, SSEA4, and TRA-1-60) in both *LBP9* and

HERVH-depleted hESC_H9. shRNA against *GFP* was used as the control (shGFP). Scale bar, 100 μ m. **g**, FACS analysis to determine the percentage of TRA-1-81⁺ cells after depletion of *LBP9* or HERVH. Three different shRNAs were employed to independently target *LBP9* and HERVH, respectively. Data are presented as mean and s.d. ($n = 3$ independent experiments with biological triplicates per experiment). **h–j**, Knockout of *LBP9* in hESCs by the CRISPR/Cas9 technology. **h**, Experimental scheme to knockout *LBP9* in hESCs using two guide RNAs (gRNAs), both targeting the second exon of *LBP9*. **i**, Analysis of *LBP9* mutant hESC clones screened by genomic PCR. **j**, Sequence analysis of the TRA-1-81 sorted cells show that *LBP9* mutants are found in differentiated (TRA-1-81⁺) but not in undifferentiated (TRA-1-81⁺) hESCs (representative samples). **k**, In contrast to human, *Tfcp2l1* (mouse *LBP9*) depletion by shRNA does not affect self-renewal (left panel) in mouse ESCs in LIF/serum condition. *Tfcp2l1*-depleted mESCs were then differentiated into embryoid bodies (right panel), and endoderm and mesoderm markers were more expressed compared with shGFP mESC-derived embryoid bodies, indicating that *Tfcp2l1*-depleted mESCs have a bias to differentiate to endoderm and mesoderm (qRT-PCR analyses). Data are normalized to *Gapdh*, and relative to shGFP expressing, undifferentiated mESCs. Error bars indicate s.d. ND indicates undetectable. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; t -test ($n = 3$ independent experiments with biological triplicates per experiment).



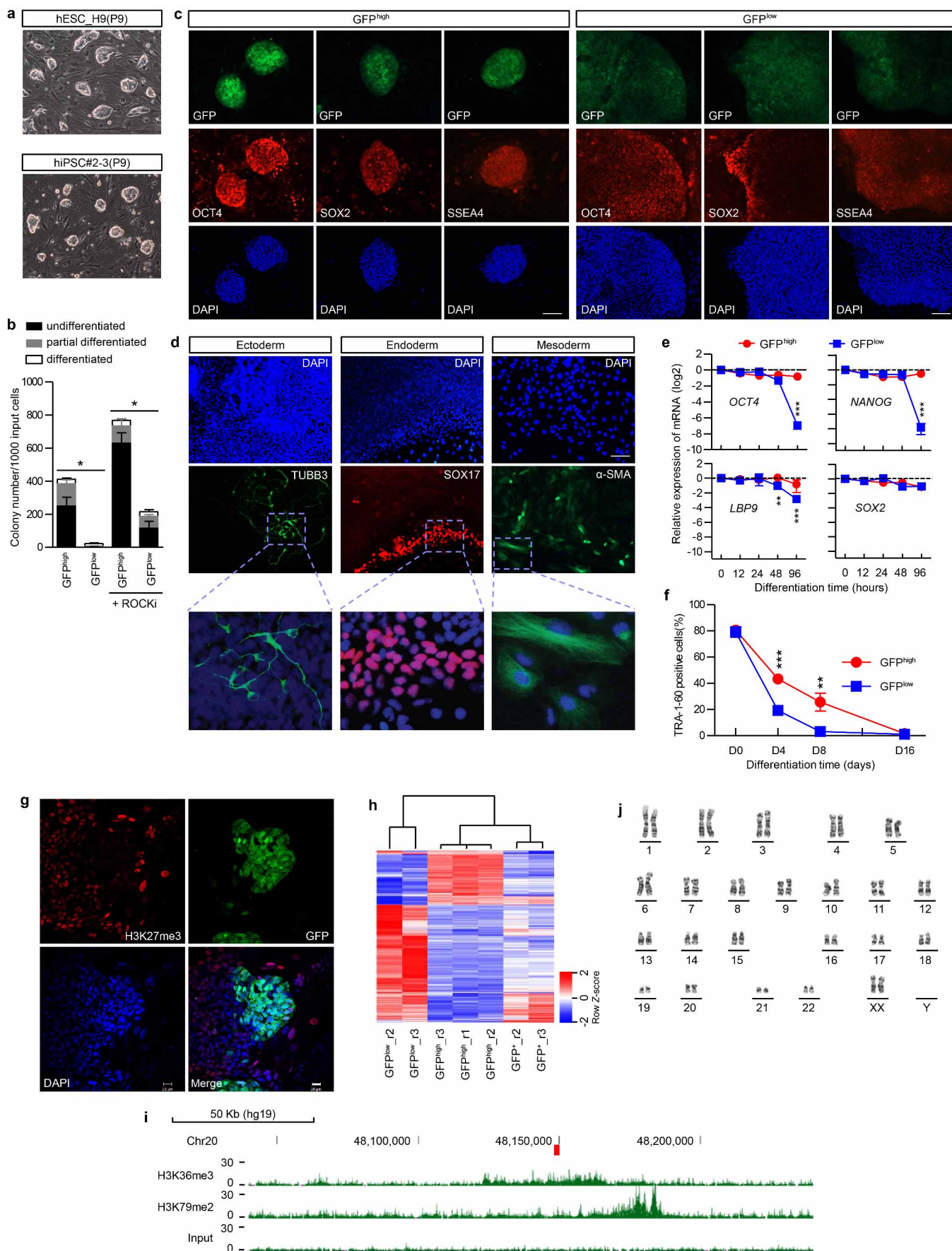
Extended Data Figure 6 | *ESRG* is required for maintenance of human pluripotency. **a**, Multi-alignment of *ESRG* putative open reading frame (ORF) from various primates. The ORF is intact in humans alone. All primate introns are shorter than the human one (which is 14,251 bp). The difference is dominantly accounted for by a single large insertion in the human sequence (around 2,000–7,500 bp) which comprises the bulk of the *ESRG* transcript (for alignment see Supplementary Data 1). **b**, Expression of *ESRG* during human embryogenesis²⁴ and in hESC cultures³ (P, passage number). **c–f**, Characterization of the effects of *ESRG* depletion on hESC_H9. Note that knockdown of *ESRG* was performed by two different shRNA constructs, #4 and #5, respectively. shRNA against GFP served as a control. **c**, *ESRG* depletion compromises hESC self-renewal, indicated by the significant decline of the expression of pluripotency markers OCT4 and SSEA4. The expression of TRA-1-60 was also decreased, while SOX2 was unaffected. The representative

images show immunostaining of pluripotency markers. Scale bar, 100 μ m. **d**, FACS analysis of TRA-1-81 expression in *ESRG*-depleted hESCs by two different shRNA constructs. Data are shown as mean and s.d. ($n = 3$ independent experiments with biological triplicates per experiment). **e**, qRT-PCR analyses of *ESRG* knockdowns using selected markers (left, pluripotency; right, differentiation). Commitment to trophoctoderm was the most apparent, characterized by the significant change in the expression of *CDX2* in the *ESRG*-depleted cells. Data, representative of three independent experiments with biological triplicates per experiment, are normalized to *GAPDH*, and relative to shGFP expressing, undifferentiated hESCs (hESC_H9). Mean and s.d.; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; t -test. **f**, Representative images of immunostaining showing expression of PAX6 (neuroectoderm) and CDX2 (trophoctoderm) in *ESRG*-depleted hESCs_H9. Scale bar, 100 μ m.



Extended Data Figure 7 | The reporter assay. **a**, Schematic of the reporter construct, pT2-LTR7-GFP#2, comprising an LTR7 region amplified from the *ESRG* locus, fused to GFP-polyA, and flanked by inverted terminal repeats (ITRs) of the SB transposon-based integration vector²². A reporter line was established by co-transfecting pT2-LTR7-GFP#2 with SB100X into HFF-1. GFP signal is detectable in both mouse and human transgenic ESCs. Representative pictures of pT2-LTR7-GFP#2-marked hESC_H9s and mESCs are shown. In the human case we show a FACS-sorted single colony. In mouse, as all cells homogeneously express GFP, we show multiple unsorted colonies. Scale bar, 100 μ m. **b**, Multiple LTR7s responding to the fibroblast-iPSC transition are capable of driving the GFP reporter. Compared to the positive control 2 (pT2-LTR7-GFP#2), four additional responsive LTR7s (#3–6) amplified from different genomic loci were tested in the reporter assay (transfected into hiPSCs). The GFP signal of the five clones correlates to the RPKM values of the RNA-seq (not shown). Mock is a negative control transfected with the empty vector (pUC19). Percentage of GFP⁺ cells (green) and mean fluorescent intensity (black) are shown. Data were obtained from three independent experiments. Error bars indicate s.d.; ** $P < 0.01$, t -test. **c**, Reporter assays to validate candidate transcription factors driving transcription from LTR7/HERVH. GFP signal is detectable in the fibroblast-derived reporter line by FACS, following forced expression of *NANOG*, *LBP9*, *OCT4*, *KLF4*, *SOX2* and *MYC* constructs. Quantification was performed at days 2 and 7 post-transfection. Control was transfected with the empty vector (pUC19). Data were obtained from two independent experiments, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; two-way ANOVA followed by Bonferroni test. A synergism between *NANOG* and *LBP9* is indicated. **d**, Schematic representation of a reporter construct (pT2-LTR7-GFP#1: wild type, WT) and its mutated version (Δ LBP9), where the LBP9 motif was deleted; the constructs were transfected into hiPSCs. FACS quantification of the GFP signal derived from wild-type and motif-deleted cells. Percentage of GFP⁺ cells (green) and mean fluorescent intensity (black) are shown. Data were obtained from three independent experiments. Error bars indicate s.d.; t -test, * $P < 0.05$.

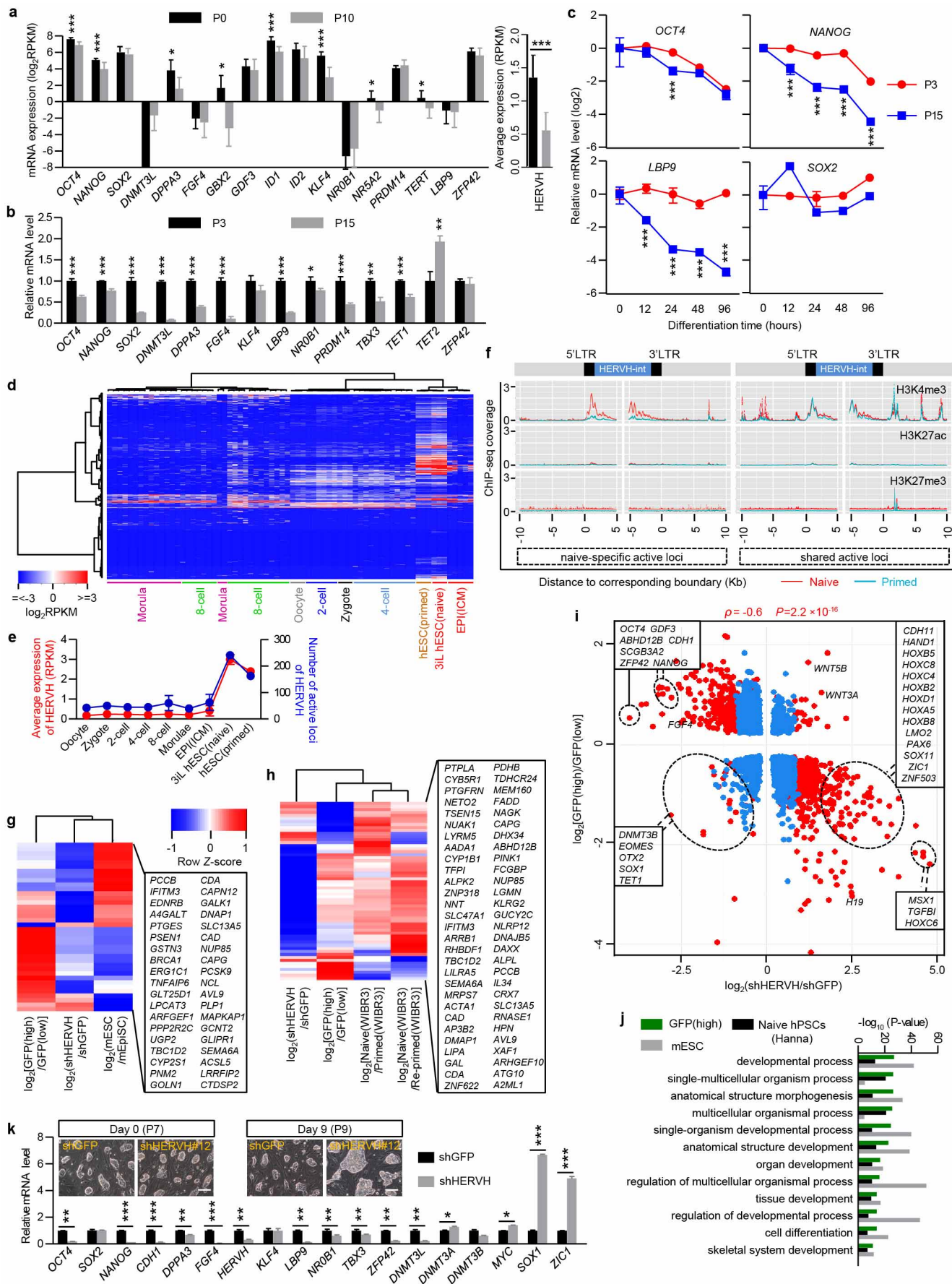
e, pT2-LTR7-GFP#2 marked, mosaic, primed hPSC colonies in conventional hESC medium consist of cells expressing HERVH at various levels, but contain GFP^{high} cell populations with mESC morphology (indicated by white arrowheads). Representative hiPSC (left panel) and hESC_H9 (right panel) colonies are shown. A GFP^{high} cell population is magnified. Scale bar, 200 μ m. **f–h**, FACS-sorted GFP^{high} and GFP^{low} hESC_H9 cells were cultured in 2i/LIF, NHSM⁴ and 3iL³ conditions, respectively. **f**, **g**, Representative images of GFP^{high} and GFP^{low} cells cultured in the different conditions at day 3. Scale bar, 200 μ m. **f**, Morphology and GFP fluorescence of GFP^{high}, 3D colonies were comparably maintained in the three different naive culture conditions, but not in primed culture conditions (KOSR and mTeSR1). **g**, Representative images show flat, GFP-negative colonies derived from GFP^{low} hESC_H9 cultured in either of the different culture conditions. **h**, Quantification by FACS of GFP-positive cells on day 6 of culturing in five media conditions: 2i/LIF, NHSM⁴, 3iL³, KOSR and mTeSR1. We cultured both GFP^{low} and GFP^{high} cells before sorting. Longer-term culturing of GFP^{high} naive cells is most compatible with 3iL³ culture condition (not shown). Percentage of GFP^{high} cells, GFP^{low} cells (bright and pale green) and mean fluorescent intensity (black) are shown. KOSR, knockout serum replacement medium. Error bars, s.d.; $n = 3$ independent cell cultures, representative of two independent experiments. **i**, **j**, Heterogeneity of GFP^{high} cells cultured in different conditions. **i**, The percentages of different hESC colonies derived from the same initial GFP^{high} population in different culture conditions. 3D/GFP^{high}, domed colony with strong GFP signal; 2D/GFP^{low}, flat colony with weak GFP signal; mosaic, colonies containing at least two cell types of GFP^{high} and either GFP^{low} or GFP⁺; 3D/GFP⁺, domed colony without detectable GFP signal; 2D/GFP⁺, flat colony without detectable GFP signal. **i**, 388–462 colonies were characterized per culture condition, using fluorescence microscopy. **j**, qRT-PCR analysis of expression levels of core pluripotency-associated transcription factors in different colony types under the 2i/LIF condition. Total RNA isolated from 10–15 colonies per colony type was reverse transcribed for qPCR. Error bars indicate s.d. ($n = 3$, technical replicates).



Extended Data Figure 8 | Characterization of LTR7-GFP-marked hPSCs.

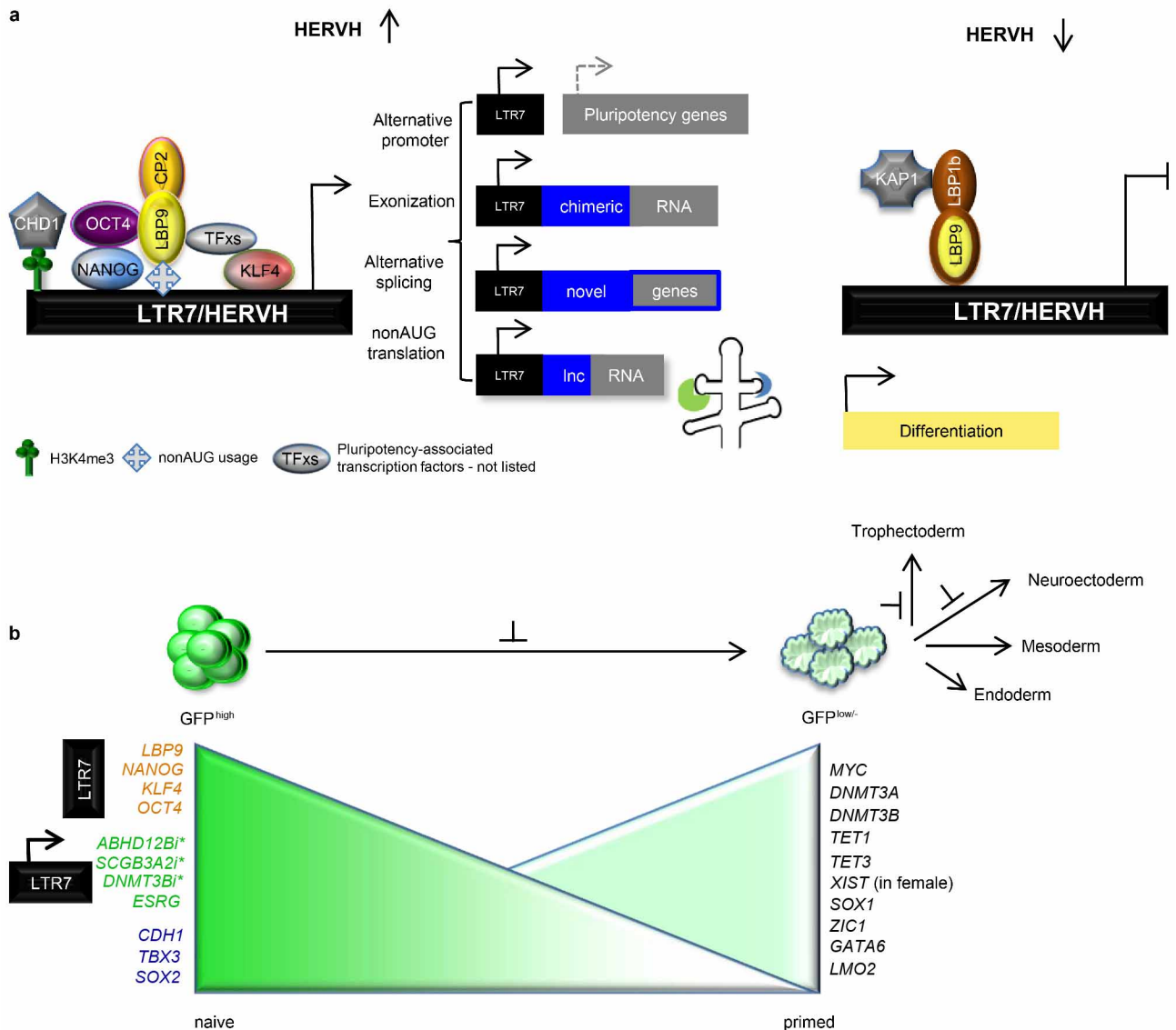
a, Genetically labelled (pT2-LTR7-GFP#2) human naive-like hESC_H9 and hiPSCs can be maintained in 2i/LIF medium for a longer period of time (followed by passage number = P9, >30 days) by re-plating (every 4–5 days), and by occasional sorting for the GFP marker. For optimal long-term culturing conditions, see Extended Data Fig. 7h. **b**, Single-cell cloning efficiency of GFP^{high} versus GFP^{low} hESC_H9. ALP-stained colonies were counted one week after plating 1,000 cells of a single cell suspension (with or without ROCK inhibitor (ROCKi)). Data were obtained from three independent experiments. Error bars indicate s.d., * $P < 0.01$, t -test. **c**, Both GFP^{high} and GFP^{low} hESCs_H9 are immunostained by the indicated pluripotency markers (OCT4, SOX2, SSEA4). Scale bar, 100 μ m. **d**, GFP^{high} cells can be differentiated, and display the markers of the three germ layers. Scale bar, 100 μ m. **e**, qRT-PCR analysis of pluripotency-associated transcription factors during *in vitro* differentiation of GFP^{high} and GFP^{low} hESC_H9s. FACS-sorted GFP^{high} and GFP^{low} cells were cultured in human 2i/LIF medium and in conventional hESC medium for 3 days, respectively, before differentiation was triggered. Error bars

indicate s.d. ($n = 3$ independent experiments with biological triplicates per experiment), ** $P < 0.01$, *** $P < 0.001$, t -test. **f**, FACS quantification of TRA-1-60-positive cells in differentiated GFP^{high} and GFP^{low} cells. Error bars indicate s.d. ($n = 3$ independent experiments with biological triplicates per experiment), t -test for each time point, ** $P < 0.01$, *** $P < 0.001$. **g**, Representative confocal image obtained after immunostaining for H3K27me3 on a chimaeric hESC_H9 colony. GFP^{high} cells (green) are marked with lower density of H3K27m3 (red) than GFP^{low} and GFP[−] cells, indicating a higher histone methylation status in the absence of GFP. Scale bar, 20 μ m. **h**, Global expression comparison between GFP^{high}, GFP⁺ and GFP^{low} cells. Hierarchical clustering of the mean expression values of global gene expression using Spearman's correlation (heat map). Biological replicates are shown. **i**, Mapping of the integration site of the pT2-LTR7-GFP#2 reporter in GFP^{high} cells. The single copy of the reporter is integrated on chromosome 20 (red box) in a transcriptionally active area, marked by H3K36me3 and H3K79me2. **j**, Karyotype analysis result indicating the normal karyotype of hESC_H9 which were used in the present study.



Extended Data Figure 9 | Transcription driven by HERVH defines naive-like state of hPSCs. **a**, Expression of pluripotency-associated transcription factors in undifferentiated early (P0) and late passage (P10) hESCs²⁴. At P10, $n = 26$; at P0, $n = 8$. t -test, $*P < 0.05$, $***P < 0.001$. **b**, qRT-PCR analysis of pluripotency-associated transcription factors in undifferentiated early (P3) and late passage (P15) hiPSCs³⁰, normalized to levels at P3. **c**, qRT-PCR analysis of pluripotency-associated transcription factors during *in vitro* differentiation of early (P3) and late passage (P15) hiPSCs. P, Passage number. t -test within each time period. **d**, Heat map showing differential HERVH transcription during human embryogenesis²⁴ and in cultured hESCs³. The raw RNA-seq data downloaded from GEO (GSE36552) and ArrayExpress (E-MTAB-2031) were analysed to monitor HERVH expression. The rows represent the expression of 1,225 full-length HERVH loci. **e**, The average transcription of and number of active HERVHs during human embryogenesis and in cultured hESCs. **f**, Chromatin status comparison around full-length HERVHs between naive and primed hESC_H1 (ref. 3). While there are no differences in shared HERVH loci, which are transcribed in both naive and primed hESCs, the 5' LTR of naive-specific HERVH loci are marked with H3K4me3. **g**, Heat map showing the comparison with mESC versus mouse epiblast stem cells (mEpiSCs⁵³) of HERVH neighbour genes. Log₂ fold change values of orthologous genes were subjected to hierarchical clustering (Pearson correlation, centroid linkage, $k = 3$). Genes selected and clustered as in **h**. **h**, The expression of neighbouring genes to HERVH in different human cell types, including GFP^{high}, HERVH-depleted hPSCs, published naive hPSCs

(naive(WIBR3)) and primed hESCs (reprimed(WIBR3))⁴. The heat map shows the comparison of row-normalized differential expression levels at log₂ scale of fold changes of GFP^{high} versus GFP^{low}, shHERVH versus shGFP, naive WIBR3 hESC versus primed and re-primed WIBR3 (GSE46872). Genes shown are those differentially expressed within every pairwise comparison (differential expression defined by log₂ modular change > 1 , with FDR cutoff at 0.01). Isoform expression merged to single gene. Samples are represented in the order of euclidean distance and were clustered using Spearman's correlation and centroid linkage. **i**, Scatter plot showing the differentially expressed genes between GFP^{high} and GFP^{low} are negatively correlated with the ones between HERVH-depleted hESCs and wild-type hESCs. The enlisted genes are enriched in GFP^{high} versus GFP^{low} and are specific to naive state (upper right), while genes downregulated by HERVH depletion are specific to primed hESCs or lineage commitment (lower). Red dots indicate differentially expressed genes. Representative clusters are shown. **j**, Gene ontology (GO) categories for downregulated genes in GFP^{high} compared to GFP^{low} as well as naive hPSCs and mESCs versus primed cells^{4,53}. **k**, Depletion of HERVH induced reduction of key transcription factors for naive hPSCs in the 2i/LIF medium. The representative images show the effects on GFP^{high} cell morphologies upon depletion of HERVH. Scale bar, 100 μ m. mRNA levels are normalized to *GAPDH*, and relative to shGFP expressing, undifferentiated hESC_H9. In **b**, **c** and **k**, error bars indicate s.d. ($n = 3$ independent experiments with biological triplicates per experiment), t -test, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.



Extended Data Figure 10 | HERVH drives a primate-specific naive pluripotency: a model. **a**, The binding sites of transcription factors for naive pluripotency are clustered on HERVH. LBP9 is a modulator of the CP2 transcription-factor family²⁸, and can form heteromeric, activator or repressor complexes with other family members, CP2, LBP1b, respectively. The activator complex interacts with OCT4 (ref. 16) and promotes pluripotency. In addition we provide evidence for the potential interaction of LBP9 and NANOG. Activated HERVHs generate numerous novel, stem-cell-specific alternative gene products. HERVH incorporates a set of regulatory lncRNAs into the network and defines novel pluripotency-associated gene products through alternative splicing (in conjunction with CHD1) or alternative non-AUG usage (in conjunction with other members of the CP2 family⁵⁵). lncRNAs, some with

a conserved domain (cruciform structure), interact with both pluripotency and chromatin modifying proteins (in green and blue). HERVH inhibits differentiation, while HERVH-derived products contribute to maintain pluripotency. LBP1b interacts with KRAB-associated protein 1 (KAP1, also called TRIM28), a repressor of ERVs during differentiation⁵⁶. **b**, GFP^{high} cells form dome-shaped colonies (3D), while GFP^{low} cells form flat (2D) colonies. Left: upregulated genes in GFP^{high} cells include (1) naive transcription factors associated with HERVH (brown); (2) LTR7/HERVH driven novel isoforms of genes (*) and novel genes (for example, *ESRG*) (green); (3) naive transcription factors shared between mice and human (blue). Right: upregulated genes in GFP^{low} cells are associated with lineage commitment.

Protein quality control at the inner nuclear membrane

Anton Khmelinskii^{1*}, Ewa Blaszczyk^{2,3*}, Marina Pantazopoulou⁴, Bernd Fischer^{5,6}, Deike J. Omnus⁴, Gaëlle Le Dez^{2,3}, Audrey Brossard^{2,3}, Alexander Gunnarsson⁴, Joseph D. Barry⁵, Matthias Meurer¹, Daniel Kirmmaier¹, Charles Boone⁷, Wolfgang Huber⁵, Gwenaël Rabut^{2,3}, Per O. Ljungdahl⁴ & Michael Knop^{1,8}

The nuclear envelope is a double membrane that separates the nucleus from the cytoplasm. The inner nuclear membrane (INM) functions in essential nuclear processes including chromatin organization and regulation of gene expression¹. The outer nuclear membrane is continuous with the endoplasmic reticulum and is the site of membrane protein synthesis. Protein homeostasis in this compartment is ensured by endoplasmic-reticulum-associated protein degradation (ERAD) pathways that in yeast involve the integral membrane E3 ubiquitin ligases Hrd1 and Doa10 operating with the E2 ubiquitin-conjugating enzymes Ubc6 and Ubc7 (refs 2, 3). However, little is known about protein quality control at the INM. Here we describe a protein degradation pathway at the INM in yeast (*Saccharomyces cerevisiae*) mediated by the Asi complex consisting of the RING domain proteins Asi1 and Asi3 (ref. 4). We report that the Asi complex functions together with the ubiquitin-conjugating enzymes Ubc6 and Ubc7 to degrade soluble and integral membrane proteins. Genetic evidence suggests that the Asi ubiquitin ligase defines a pathway distinct from, but complementary to, ERAD. Using unbiased screening with a novel genome-wide yeast library based on a tandem fluorescent protein timer⁵, we identify more than 50 substrates of the Asi, Hrd1 and Doa10 E3 ubiquitin ligases. We show that the Asi ubiquitin ligase is involved in degradation of mislocalized integral membrane proteins, thus acting to maintain and safeguard the identity of the INM.

To identify components of INM quality control, we focused on the ubiquitin-conjugating enzyme Ubc6. Ubc6 is an integral membrane protein that localizes to the endoplasmic reticulum and the INM where it targets for degradation soluble and integral membrane proteins together with Ubc7 and Doa10 (refs 6, 7). We established a microscopy-based bimolecular fluorescence complementation (BiFC) assay⁸ to screen for new E3 ubiquitin ligases interacting with Ubc6 (Fig. 1a). In total, 10 out of 54 known or putative E3s, including Doa10, interacted with Ubc6 at distinct subcellular locations (Fig. 1b and Extended Data Fig. 1a). Among these, Asi1 and Asi3 displayed a BiFC signal restricted to the nuclear rim (Fig. 1b). Despite their colocalization at the endoplasmic reticulum, no interaction was detected between Ubc6 and Hrd1 (Extended Data Fig. 1a), suggesting a low rate of false-positive interactions in our BiFC assay.

Asi1 and Asi3 are integral membrane RING domain proteins of the INM and form the Asi complex^{4,9,10}. Together with the INM protein Asi2, the Asi complex functions in the Ssy1-Ptr3-Ssy5 (SPS) amino-acid-sensing pathway, where it is involved in the degradation of Stp1 and Stp2 transcription factors¹¹. We tested the interactions of Asi1 and Asi3 with all E2 ubiquitin-conjugating enzymes using the BiFC assay. In addition to Ubc6, Asi1 and Asi3 interacted with Ubc7 and weakly

with Ubc4 (Extended Data Fig. 1b–d). We validated these interactions in microscale thermophoresis experiments¹² with recombinant proteins (Fig. 1c and Extended Data Fig. 1e). The Ubc7-binding region of Cue1 (Cue1^{U7BR})¹³, a protein that tethers Ubc7 to the endoplasmic reticulum membrane¹⁴, was included in the assays. A carboxy-terminal fragment of Hrd1 (Hrd1^{CT}) expected to interact with Ubc7 but not Ubc6 served as control^{2,3}. The RING domains of Asi1 and Asi3 (Asi1^{RING} and Asi3^{RING}) interacted with Ubc7, provided it was bound to Cue1^{U7BR}, with affinities similar to Hrd1^{CT}. Asi1^{RING} and Asi3^{RING}, but not Hrd1^{CT}, also interacted weakly with Ubc6 lacking its transmembrane domain (Ubc6^{ΔTM}) (Fig. 1c).

The Asi proteins maintain the SPS pathway in the ‘off state’ in the absence of inducing amino acids, and do so by targeting for proteasomal degradation the low levels of Stp1 and Stp2 that inadvertently mislocalize into the nucleus¹¹. Consequently, *asi* mutants exhibit aberrant constitutive Stp1/Stp2-dependent transcription⁹. We observed that *ubc7Δ* and, to a lesser extent, *ubc6Δ* mutants exhibited increased expression of Stp1/Stp2-regulated genes similar to the *asi1Δ* and *asi3Δ* mutants (Fig. 1d and Extended Data Fig. 1f). These effects were not due to inactivation of Hrd1 or Doa10 ubiquitin ligases (Extended Data Fig. 1f), thus implicating Ubc6 and Ubc7 in the SPS pathway.

Next, we assayed the ubiquitylation of an artificial Asi substrate based on the first 45 amino acids of Stp2 (Stp2^N). This fragment of Stp2 contains a degron that is recognized by the Asi complex¹¹. Ubiquitylation of Stp2^N fused to the tandem affinity purification (TAP) tag was reduced in *ubc6Δ* and severely impaired in *asi3Δ* and *ubc7Δ* mutants (Fig. 1e). In addition, ubiquitylation of Stp1 and Stp2 mutants with constitutive SPS-independent nuclear localization was impaired in *asi1Δ* and *asi3Δ* strains (Extended Data Fig. 1g). Together, these results establish the Asi complex as an E3 ubiquitin ligase of the INM that functions with Ubc6 and Ubc7.

Functionally related genes can be identified by similarity of genetic interaction profiles¹⁵. We searched for novel functions of the Asi ubiquitin ligase by mining a genome-scale genetic interaction map¹⁶. In this data set, the fitness of 5.4 million double-mutant combinations was measured by colony size, generating genetic interaction profiles for ~75% of all *S. cerevisiae* genes. We calculated correlation coefficients between genetic interaction profiles of *ASI* genes and the other 4,458 genes in the genetic interaction map. In this analysis, the genetic interaction profiles of *ASI* genes correlated with each other and, to a similar extent, with *HRD1*, *DOA10*, *UBC6*, *UBC7* and *CUE1* among others (Fig. 2a and Supplementary Table 1), suggesting that Asi and ERAD E3 ubiquitin ligases are functionally related. We sought to determine whether they work in the same or parallel pathways. Strains lacking *HRD1* and

¹Zentrum für Molekulare Biologie der Universität Heidelberg (ZMBH), DKFZ-ZMBH Alliance, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany. ²Centre National de la Recherche Scientifique, UMR 6290, 35000 Rennes, France. ³Institut de Génétique et Développement de Rennes, Université de Rennes 1, 35000 Rennes, France. ⁴Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Svante Arrhenius väg 20B, SE-106 91 Stockholm, Sweden. ⁵Genome Biology Unit, European Molecular Biology Laboratory (EMBL), Meyerhofstraße 1, 69117 Heidelberg, Germany. ⁶Computational Genome Biology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 580, 69120 Heidelberg, Germany. ⁷Department of Molecular Genetics, Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, 160 College St, Toronto, Ontario M5S 3E1, Canada. ⁸Cell Morphogenesis and Signal Transduction, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany.

*These authors contributed equally to this work.

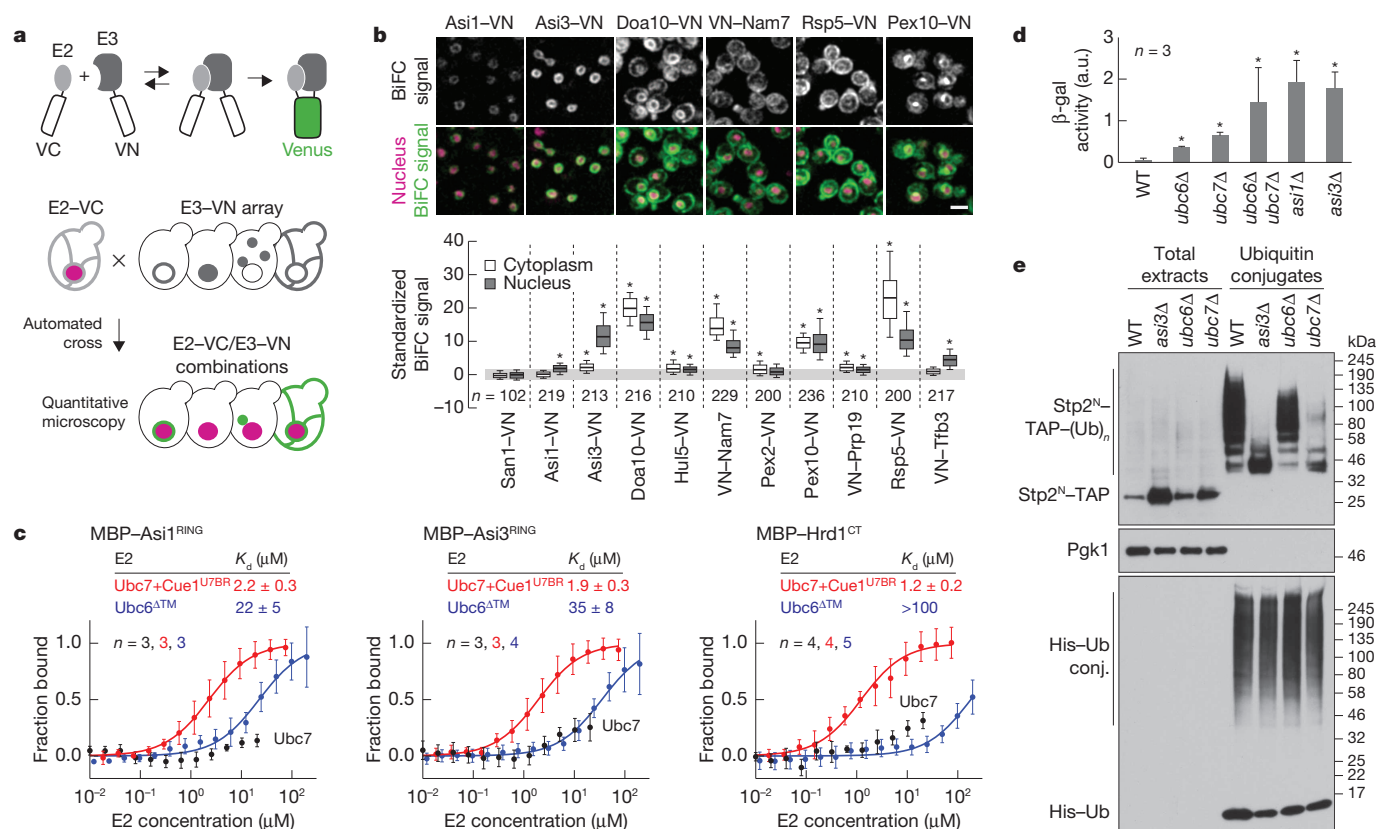


Figure 1 | The Asi complex is a Ubc6/Ubc7-dependent E3 ubiquitin ligase of the INM. **a**, BiFC strategy used to assay E2–E3 interactions. E2 and E3 proteins were endogenously tagged with carboxy- and amino-terminal fragments of the Venus fluorescent protein (VC and VN, respectively). Interactions between E2 and E3 proteins enable reconstitution of functional Venus that is detected with fluorescence microscopy. Rpn7 fused to the red fluorescent protein tDimer2 served as a nuclear marker. **b**, Quantification of BiFC signals in cells co-expressing VC–Ubc6 and VN-tagged E3s. Fluorescence microscopy examples representative of six fields of view (top). Scale bar, 5 μm. BiFC signals were measured in the cytoplasm and nucleus of individual cells (bottom, *n* as shown). Whiskers extend from the tenth to ninetieth percentiles. **c**, Microscale thermophoresis analysis of interactions between recombinant maltose binding protein (MBP)–E3 fragments and the indicated E2s. Plots show the

fraction of MBP–E3 bound to the E2 at each tested E2 concentration (mean ± s.d., *n* as shown). Dissociation constants (K_d , mean ± s.d.) were derived from nonlinear fits with the law of mass action (solid lines). **d**, Activity of β-galactosidase (β-gal) expressed from the *AGP1* promoter in the indicated strains (mean ± s.d., *n* = 3 clones). a.u., arbitrary units; WT, wild type. **e**, Ubiquitylation of Stp2^N-TAP in strains expressing 10×histidine (His)-tagged ubiquitin. Total cell extracts and ubiquitin conjugates eluted after immobilized-metal affinity chromatography were separated by SDS-PAGE followed by immunoblotting with antibodies against the TAP tag, Pgl1 and ubiquitin. Representative immunoblots from three technical replicates. * $P < 10^{-4}$ (**b**; one-way analysis of variance (ANOVA) with Bonferroni correction for multiple testing) and * $P < 0.05$ (**d**; two-tailed *t*-test).

the unfolded protein response genes *IRE1* or *HAC1* show impaired growth at increased temperature¹⁷. Additional deletion of *ASI1* resulted in a synthetic lethal phenotype under these conditions¹⁸ (Fig. 2b and Extended Data Fig. 2), suggesting that Asi1 and Hrd1 function in parallel pathways.

We used a tandem fluorescent protein timer (tFT) approach⁵ to perform unbiased proteome-wide screens for substrates of the Asi, Hrd1 and Doa10 ubiquitin ligases. A tFT is a tag composed of two fluorescent proteins (mCherry and superfolder green fluorescent protein (sfGFP)) with distinct fluorophore maturation rates. The mCherry/sfGFP intensity ratio is a measure of protein degradation kinetics in steady state (Fig. 3a), with a dynamic range and sensitivity that exceed conventional cycloheximide chase experiments⁵ (Supplementary Note 1). We constructed a genome-wide library of yeast strains each expressing a different tFT-tagged protein (Supplementary Methods). Library construction relied on a seamless tagging strategy that minimizes the influence of the tag on gene expression¹⁹ (Extended Data Fig. 3a). In total, 4,044 proteins were successfully tagged to create a tFT library covering ~73% of verified or uncharacterized open reading frames in the *S. cerevisiae* genome (Supplementary Table 2). We introduced *asi1Δ*, *asi3Δ*, *hrd1Δ*, *doa10Δ*, *ubc6Δ* and *ubc7Δ* deletion alleles into the tFT library using high-throughput genetic crosses²⁰. The effect of each gene deletion on the stability of each protein in the library was examined with high-throughput fluorescence measurements of colonies⁵ (Extended Data Fig. 3b) and quantified as a

z-score. More proteins were stabilized (positive *z*-score) than destabilized in the six mutants (Extended Data Fig. 3c and Supplementary Table 3), in agreement with the role of Asi, Hrd1 and Doa10 ubiquitin ligases in protein degradation. Hierarchical clustering of top hits recapitulated known E2–E3 interactions and revealed three clusters of 20, 30 and 9 potential substrates for the Asi, Hrd1 and Doa10 ubiquitin ligases, respectively (Fig. 3b). Hrd1 substrates, including the known substrate Der1 (ref. 21), were stabilized only in the *ubc7Δ* mutant, whereas Doa10 substrates were stabilized in both *ubc6Δ* and *ubc7Δ* mutants. Most Asi substrates, including the recently identified Erg11 (ref. 18), were stabilized in the *ubc7Δ* mutant with only weak effects of the *ubc6Δ* mutant (Fig. 3b). Stp1 and Stp2 were not identified as Asi substrates in the screen, probably the consequence of their efficient targeting for degradation by the E3 ubiquitin ligase SCF^{Grr1} in the cytoplasm¹¹. The vast majority of potential substrates in each set were integral membrane or secretory proteins distributed along the endomembrane system and the Hrd1 and Asi substrates were enriched in endoplasmic reticulum and vacuolar proteins (Fig. 3c, d and Extended Data Fig. 3d, e). These findings are consistent with the organization and functions of endoplasmic-reticulum-associated ubiquitin ligases, thus establishing the tFT library as a valuable resource for studies of protein degradation (Supplementary Note 2), and indicate that the Asi complex is involved in degradation of a distinct set of integral membrane proteins.

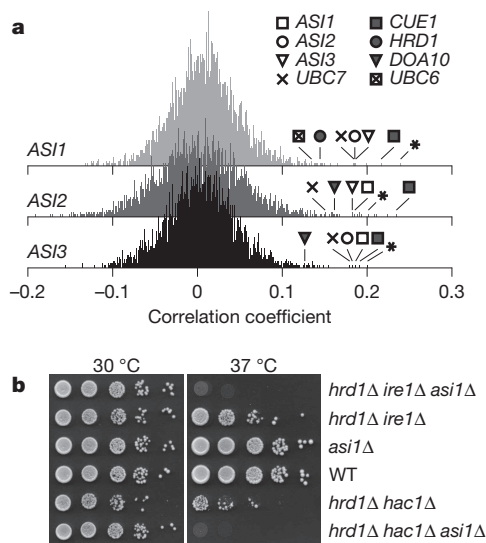


Figure 2 | Functional overlap between Asi and ERAD E3 ubiquitin ligases.
a, Histograms of Pearson correlation coefficients calculated between the genetic interaction profiles of each ASI gene and ~75% of all yeast genes, obtained from a previously published genome-scale genetic interaction map¹⁶. Asterisks mark the dubious open reading frame YMR119W-A, which overlaps with the ASI1 gene. **b**, Tenfold serial dilutions of strains grown on synthetic complete medium for 2 days at 30 or 37 °C.

Several Asi substrates that were reproducibly stabilized in *asi1Δ* and *asi3Δ* mutants were not stabilized in strains lacking ASI2 (Extended Data Fig. 4a), suggesting that Asi2 might function as a substrate-specific recognition factor. The Asi2-independent nature of the interaction between Asi3 and Ubc6 further supports this notion (Extended Data Fig. 4b). With the exception of Aqy2, which was not expressed during exponential growth in liquid medium, all tFT-tagged substrates localized to the endoplasmic reticulum in wild-type cells and eight of them accumulated at the nuclear rim specifically in the *asi1Δ* mutant (Fig. 3e and Extended Data Fig. 4c). This result is consistent with protein stabilization at the INM where the Asi proteins reside. Cycloheximide chase experiments with haemagglutinin epitope (HA)-tagged variants revealed substantial turnover of Vtc1, Erg11, Vcx1 and Vtc4 in wild-type cells. All four proteins were stabilized specifically in the absence of ASI1 (Fig. 3f and Extended Data Fig. 5), further validating our screening approach (Supplementary Note 1). Interestingly, Vtc1 and Vtc4 were previously shown to localize to the vacuolar membrane²². Both

We analysed this novel function of the Asi ubiquitin ligase with ten tFT-tagged substrates. Genetic crosses with additional deletion mutants revealed the involvement of Cue1 in Asi-dependent degradation (Extended Data Fig. 4a), in agreement with our biochemical analysis (Fig. 1c).

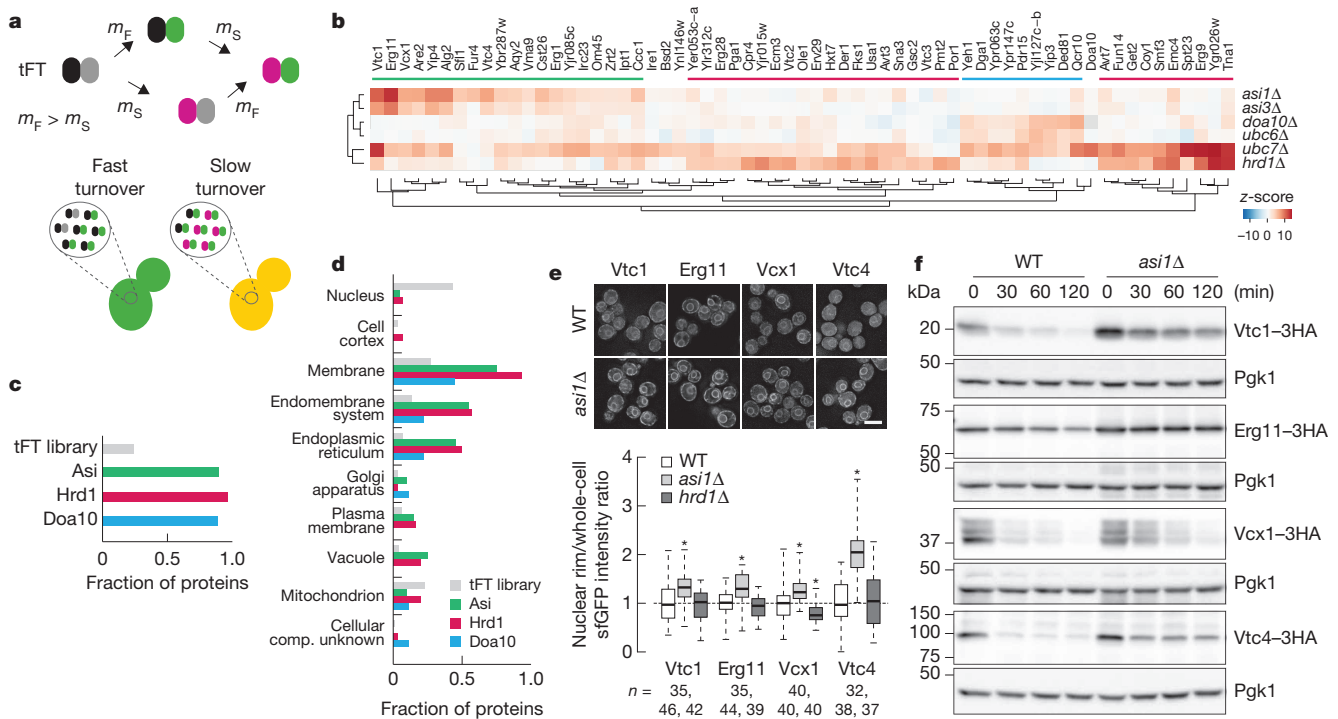


Figure 3 | Systematic identification of substrates for Asi and ERAD E3 ubiquitin ligases. **a**, A tandem fluorescent protein timer (tFT) is composed of two fluorescent proteins: one more slowly maturing (for example, the red fluorescent protein mCherry, maturation rate constant m_S) and the other faster maturing (for example, the green fluorescent protein sfGFP, maturation rate constant m_F). When fused to a protein of interest, a tFT reports on the degradation kinetics of the fusion protein: whereas fusions undergoing fast turnover are degraded before mCherry maturation, resulting in a low mCherry/sfGFP intensity ratio, the relative fraction of mature mCherry increases for proteins with slower turnover. **b**, Summary heat map of the screens for tFT-tagged proteins with altered stability in the indicated mutants. Changes in protein stability (z-score) are colour-coded from blue (decrease) to red (increase). Only proteins with a significant change in stability in at least one mutant (1% false discovery rate and z-score > 4) are shown. Clusters of potential substrates of Asi (green), Hrd1 (red) and Doa10 (blue) E3 ubiquitin ligases are indicated. **c**, **d**, Fraction of proteins in the tFT library and in the three

clusters in **b** with a predicted transmembrane domain or signal peptide (**c**) or mapped to component Gene Ontology (GO) terms (**d**). Each cluster is significantly enriched in proteins with a predicted transmembrane domain or signal peptide compared to the tFT library ($P < 2.2 \times 10^{-16}$, Fisher's exact test). **e**, Quantification of sGFP signals in strains expressing tFT-tagged proteins from the Asi cluster in **b**. Fluorescence microscopy examples representative of five fields of view (top). Scale bar, 5 μm. sGFP intensities were measured in individual cells and at the nuclear rim (bottom, n as shown). For each protein, measurements were normalized to the mean of the respective wild type. Whiskers extend from minimum to maximum values. * $P < 0.05$ (two-tailed t -test). **f**, Degradation of 3×HA-tagged proteins after blocking translation with cycloheximide. Whole-cell extracts were separated by SDS-PAGE followed by immunoblotting with antibodies against the HA tag and Pgk1 as loading control. Representative immunoblots from three technical replicates.

proteins mislocalize to the endoplasmic reticulum and nuclear rim only on overexpression or C-terminal tagging (Extended Data Fig. 6). Whether the Asi ubiquitin ligase recognizes such mislocalized proteins through specific degrons, as is the case with Stp1 and Stp2 transcription factors¹¹, or other features such as compartment-specific properties of transmembrane domains²³ is an open question.

The nuclear pore complex establishes a barrier between the cytoplasm and the nucleoplasm. However, increasing evidence suggests that not only small soluble proteins but also integral membrane proteins with cytoplasmic domains of up to 60 kilodaltons (kDa) can passively diffuse past the nuclear pore, the latter through a ~10 nm side channel^{6,24–28}. We propose that the Asi ubiquitin ligase targets such mislocalized and potentially harmful proteins for degradation. Although the Asi proteins are not obviously conserved outside of yeast, the general importance of membrane-associated protein degradation mechanisms and the large diversity of integral membrane RING domain proteins in mammalian cells²⁹ suggest that dedicated E3 ubiquitin ligases functioning in INM-associated protein degradation exist also in metazoans.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 21 September; accepted 20 November 2014.

- Mekhalil, K. & Moazed, D. The nuclear envelope in genome organization, expression and stability. *Nature Rev. Mol. Cell Biol.* **11**, 317–328 (2010).
- Zattas, D. & Hochstrasser, M. Ubiquitin-dependent protein degradation at the yeast endoplasmic reticulum and nuclear envelope. *Crit. Rev. Biochem. Mol. Biol.* <http://dx.doi.org/10.3109/10409238.2014.959889> (2014).
- Ruggiano, A., Foresti, O. & Carvalho, P. Quality control: ER-associated degradation: protein quality control and beyond. *J. Cell Biol.* **204**, 869–879 (2014).
- Zargari, A. *et al.* Inner nuclear membrane proteins Asi1, Asi2, and Asi3 function in concert to maintain the latent properties of transcription factors Stp1 and Stp2. *J. Biol. Chem.* **282**, 594–605 (2007).
- Khmelnikii, A. *et al.* Tandem fluorescent protein timers for *in vivo* analysis of protein dynamics. *Nature Biotechnol.* **30**, 708–714 (2012).
- Deng, M. & Hochstrasser, M. Spatially regulated ubiquitin ligation by an ER/nuclear membrane ligase. *Nature* **443**, 827–831 (2006).
- Boban, M., Pantazopoulou, M., Schick, A., Ljungdahl, P. O. & Foisner, R. A nuclear ubiquitin-proteasome pathway targets the inner nuclear membrane protein Asi2 for degradation. *J. Cell Sci.* **127**, 3603–3613 (2014).
- Hu, C.-D., Chinenov, Y. & Kerppola, T. K. Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol. Cell* **9**, 789–798 (2002).
- Forsberg, H., Hammar, M., Andréasson, C., Molinér, A. & Ljungdahl, P. O. Suppressors of *ssy1* and *ptr3* null mutations define novel amino acid sensor-independent genes in *Saccharomyces cerevisiae*. *Genetics* **158**, 973–988 (2001).
- Boban, M. *et al.* Asi1 is an inner nuclear membrane protein that restricts promoter access of two latent transcription factors. *J. Cell Biol.* **173**, 695–707 (2006).
- Omnus, D. J. & Ljungdahl, P. O. Latency of transcription factor Stp1 depends on a modular regulatory motif that functions as cytoplasmic retention determinant and nuclear degron. *Mol. Biol. Cell* **25**, 3823–3833 (2014).
- Wienken, C. J., Baaske, P., Rothbauer, U., Braun, D. & Duhr, S. Protein-binding assays in biological liquids using microscale thermophoresis. *Nature Commun.* **1**, 100 (2010).
- Kostova, Z., Mariano, J., Scholz, S., Koenig, C. & Weissman, A. M. A Ubc7p-binding domain in Cue1p activates ER-associated protein degradation. *J. Cell Sci.* **122**, 1374–1381 (2009).
- Biederer, T., Volkwein, C. & Sommer, T. Role of Cue1p in ubiquitination and degradation at the ER surface. *Science* **278**, 1806–1809 (1997).
- Costanzo, M., Baryshnikova, A., Myers, C. L., Andrews, B. & Boone, C. Charting the genetic interaction map of a cell. *Curr. Opin. Biotechnol.* **22**, 66–74 (2011).
- Costanzo, M. *et al.* The genetic landscape of a cell. *Science* **327**, 425–431 (2010).
- Friedlander, R., Jarosch, E., Urban, J., Volkwein, C. & Sommer, T. A regulatory link between ER-associated protein degradation and the unfolded-protein response. *Nature Cell Biol.* **2**, 379–384 (2000).
- Foresti, O., Rodriguez-Vaello, V., Funaya, C. & Carvalho, P. Quality control of inner nuclear membrane proteins by the Asi complex. *Science* **346**, 751–755 (2014).
- Khmelnikii, A., Meurer, M., Duishoev, N., Delhomme, N. & Knop, M. Seamless gene tagging by endonuclease-driven homologous recombination. *PLoS ONE* **6**, e23794 (2011).
- Baryshnikova, A. *et al.* Synthetic genetic array (SGA) analysis in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Methods Enzymol.* **470**, 145–179 (2010).
- Zattas, D., Adle, D. J., Rubenstein, E. M. & Hochstrasser, M. N-terminal acetylation of the yeast Derlin Der1 is essential for Hrd1 ubiquitin-ligase activity toward luminal ER substrates. *Mol. Biol. Cell* **24**, 890–900 (2013).
- Uttenweiler, A., Schwarz, H., Neumann, H. & Mayer, A. The vacuolar transporter chaperone (VTC) complex is required for microautophagy. *Mol. Biol. Cell* **18**, 166–175 (2007).
- Sharpe, H. J., Stevens, T. J. & Munro, S. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell* **142**, 158–169 (2010).
- Meinema, A. C., Poolman, B. & Veenhoff, L. M. The transport of integral membrane proteins across the nuclear pore complex. *Nucleus* **3**, 322–329 (2012).
- Ellenberg, J. *et al.* Nuclear membrane dynamics and reassembly in living cells: targeting of an inner nuclear membrane protein in interphase and mitosis. *J. Cell Biol.* **138**, 1193–1206 (1997).
- Soullam, B. & Worman, H. J. The amino-terminal domain of the lamin B receptor is a nuclear envelope targeting signal. *J. Cell Biol.* **120**, 1093–1100 (1993).
- Hinshaw, J. E., Carragher, B. O. & Milligan, R. A. Architecture and design of the nuclear pore complex. *Cell* **69**, 1133–1141 (1992).
- Beck, M., Lucić, V., Förster, F., Baumeister, W. & Medalia, O. Snapshots of nuclear pore complexes in action captured by cryo-electron tomography. *Nature* **449**, 611–615 (2007).
- Nakamura, N. The Role of the transmembrane RING finger proteins in cellular and organelle function. *Membranes* **1**, 354–393 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. Lemberg, E. Schiebel and B. Bukau for support and discussions, A. Kaufmann, C.-T. Ho, A. Bartosik and B. Besenbeck for help with tFT library construction, K. Ryman for the qRT-PCR analysis of gene expression, M. Hochstrasser for strains, the GeneCore and the media kitchen facilities of the European Molecular Biology Laboratory (EMBL) and Donnelly Centre for support with infrastructure and media. This work was supported by the Sonderforschungsbereich 1036 (SFB1036, TP10) from the Deutsche Forschungsgemeinschaft (DFG) (M.K.), the Swedish Research Council grant VR2011-5925 (P.O.L.), INSERM and grants from ANR (ANR-12-JSV8-0003-001) and Biosit (G.R.), fellowships from the European Molecular Biology Organization (EMBO ALTF 1124-2010 and EMBO ASTF 546-2012) (A.K.) and fellowships from the Ministère de la Recherche et de l'Enseignement Supérieur and La Ligue Contre le Cancer (E.B.). M.K. received funds from the CellNetworks Cluster of Excellence (DFG) for support with tFT library construction. W.H. acknowledges funding from the EC Network of Excellence Systems Microscopy. C.B. was supported by funds from the Canadian Institute for Advanced Research (GNE-BOON-141871), National Institutes of Health (R01HG005853-01), Canadian Institute for Health Research (MOP-102629) and the National Science and Engineering Research Council (RGPIN 204899-6).

Author Contributions G.R. designed the BiFC, microscale thermophoresis and ubiquitin pull-down experiments that were performed by E.B., G.L.D. and A.B. M.P., D.J.O. and A.G. contributed the biochemical analysis of Asi-dependent ubiquitylation. M.K. and A.K. designed and coordinated the tFT project. A.K. and M.M. designed and constructed the tFT library and performed the screens with help from D.K. and C.B. B.F. and J.D.B. developed the screen analysis methods, with input from A.K., M.K., W.H. and C.B. M.K., A.K., G.R. and P.O.L. prepared the figures and wrote the paper with input from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.K. (m.knop@zmbh.uni-heidelberg.de), P.O.L. (per.ljungdahl@su.se) or G.R. (gwenael.rabut@inserm.fr).

METHODS

Yeast methods and plasmids. Yeast genome manipulations (gene deletions and tagging) were performed using PCR targeting, as described³⁰. Yeast strains and plasmids used in this study are listed in Supplementary Tables 4 and 5, respectively.

β -galactosidase activity assay. Cells were grown in synthetic minimal medium and β -galactosidase activity was measured in *N*-lauroyl-sarcosine-permeabilized cells as described³¹.

RNA isolation and qRT-PCR. Strains with auxotrophies complemented by plasmids pRS316 (*URA3*), pRS317 (*LYS2*) and pAB1 (*HIS3*, *MET15* and *LEU2*) were grown in synthetic minimal medium to 10^7 cells ml⁻¹ and collected by centrifugation. RNA was isolated using the RiboPure Yeast Kit and treated with Turbo-DNase (Ambion). The quality of RNA preparations was assessed by electrophoresis on a 1% agarose gel with 10 mM guanidine thiocyanate, and the lack of DNA contamination was confirmed by PCR. One microgram of RNA was used for complementary DNA synthesis with oligo (dT)12-19 (Invitrogen) using SuperScript III Reverse Transcriptase (Life Technologies). Quantitative reverse transcriptase PCR (qRT-PCR) reactions were prepared using Kapa SybrFast qPCR Master Mix (Kapa-Biosystems). cDNA mixtures were diluted 1:40 and 5 μ l were used in a reaction volume of 20 μ l with the following primer pairs: *AGP1* fwd 5'-CTGCCGTGCGTAGGTTTT-3' and *AGP1* rev 5'-AGAAGAAGGTGAGATAGCCGA-3'; *GNP1* fwd 5'-CACCACAAGAACAAGAACAGAAAC-3' and *GNP1* rev 5'-ACCGACCAGCAAACAGTA-3'; *TAF10* fwd 5'-ATATTCCAGGATCAGGTCTTCCGTAGC-3' and *TAF10* rev 5'-CAACAACAACATCAACAGAATGAGAAGACTAC-3'.

The levels of gene expression in three biological replicates were determined in two separate amplifications with triplicate technical replicates of each of the three genes analysed using the comparative ΔC_T method (RotorGene 6000, Corbett Life Science). Relative levels of *AGP1* and *GNP1* messenger RNA were normalized with respect to the levels of the invariant reference gene *TAF10*; the levels of *AGP1* and *GNP1* in strains carrying the indicated mutations were subsequently averaged and normalized to the levels of expression in the corresponding isogenic wild-type strains.

Purification of decahistidine-ubiquitin protein conjugates. Ubiquitylated proteins were purified from 1×10^9 exponentially growing yeast cells expressing 10 \times His-tagged ubiquitin using a protocol adapted from ref. 32. Cell pellets were resuspended in 2 ml 20% trichloroacetic acid and lysed for 2 min using glass beads in a Disrupter Genie homogenizer (Scientific Industries). After precipitation, proteins were resuspended in 3 ml guanidium buffer (6 M guanidinium chloride, 100 mM Tris-HCl, pH 9, 300 mM NaCl, 10 mM imidazole, 0.2% Triton X-100 and 5 mM chloroacetamide), clarified at 30,000g and incubated for 1.5 h at room temperature with TALON Metal Affinity Resin (Clontech). The beads were then washed with wash buffer (8 M urea, 100 mM sodium phosphate, pH 7.0, 300 mM NaCl, 5 mM imidazole, 0.2% Triton X-100 and 5 mM chloroacetamide) containing 0.2% SDS (twice) and lacking SDS (twice). 10 \times His-ubiquitin conjugates were finally eluted with 200 μ l elution buffer (8 M urea, 100 mM sodium phosphate, pH 7.0, 300 mM NaCl, 250 mM imidazole, 0.2% Triton X-100 and 5 mM chloroacetamide). Total extracts (1% of the amount used for purification) and ubiquitin conjugate eluates were analysed by SDS-PAGE and immunoblotting with antibodies against the TAP tag (PAP, 1:1,000, Sigma). As controls, levels of ubiquitin conjugates and Pgk1 were assessed with anti-ubiquitin (P4D1 horseradish peroxidase (HRP) conjugate, 1:1,000, Santa Cruz) and anti-Pgk1 antibodies (clone 22C5D8, 1:10,000, Invitrogen), respectively. Immunogenic proteins were detected by chemiluminescence using SuperSignal West Femto Substrate (Thermo Scientific) and recorded using autoradiographic films (CP-BU, Agfa) processed with a Curix 60 developing machine (Agfa).

Purification of hexahistidine-ubiquitin protein conjugates. Ubiquitylated proteins were purified from 5×10^8 exponentially growing yeast cells expressing 6 \times His-tagged ubiquitin as previously described³³. 6 \times His-ubiquitin conjugates were retained on nickel-nitrilotriacetic acid Sepharose beads (Qiagen) and eluted in the presence of 300 mM (Stp1-HA, Stp1-RI₁₇₋₃₃-HA) or 500 mM (Stp2-HA, Stp2 Δ_{2-13} -HA) imidazole. Total extracts, flow-through and eluate fractions were precipitated with 10% trichloroacetic acid, analysed by SDS-PAGE and immunoblotting with antibodies against the haemagglutinin tag (1:5,000, Roche) and the signals were recorded using autoradiographic film (CL-Xposure, Thermo Scientific). As controls, levels of ubiquitin conjugates and Pgk1 were assessed with anti-His₅ (1:5,000, Qiagen) and anti-Pgk1 antibodies (1:10,000, Invitrogen), respectively, and detected by chemiluminescence using SuperSignal West Dura Extended Duration Substrate (Thermo Scientific) and a Molecular Imager ChemiDoc XRS+ with Image Lab v3 build 11 software (BioRad). Loaded total and flow-through fractions correspond to 2% (Stp1-HA or Stp1-RI₁₇₋₃₃-HA) and 0.7% (Stp2-HA or Stp2 Δ_{2-13} -HA) of the amount used for purification of ubiquitin conjugates.

Bimolecular fluorescence complementation. BiFC interaction assays were performed using E2 and E3 proteins tagged with the VC173 and VN155 fragments (VC and VN) of the Venus fluorescent protein, respectively³⁴. All E2 and E3 proteins

were tagged C-terminally, with the following exceptions that were N-terminally tagged: Ubc6, because the C terminus of Ubc6 faces the endoplasmic reticulum lumen³⁵; Ubc7, to preserve its interaction with Cue1 (ref. 36); Ubc1, because the growth of strains expressing Ubc1 endogenously tagged at the C terminus with VC appears compromised; the E3 proteins Far1, Mot2, Nam7, Prp19, Ste5 and Tfb3, as they all have their E2 binding domain at the N terminus. All fusions were expressed from their endogenous chromosomal loci, with the exception of Rsp5-VN, which was expressed from its endogenous promoter on the centromeric plasmid pGR703 (Supplementary Table 5).

Strains expressing VC-tagged E2 proteins were constructed in the scEB115 background. scEB115 carries markers for selection of haploid progeny in automated crosses (*can1::STE2pr-spHIS5* and *lyp1::STE3pr-HPH*) and expresses the proteasomal subunit Rpn7 fused to the red fluorescent protein tDimer2 as nuclear marker (Supplementary Table 4). Strains expressing VN-tagged E3 proteins were either obtained from a commercially available collection (Bioneer Corporation) or constructed by homologous recombination in the BY4741 background. Expression of VC- and TAP-tagged fusions was validated by immunoblotting with mouse anti-GFP (clones 7.1 and 13.1, Roche) and peroxidase anti-peroxidase (Sigma) antibodies to detect the VC and TAP tag, respectively, and mouse anti-actin (clone c4, Merck Millipore) for loading controls.

Strains expressing individual E2 and E3 protein fusions were crossed to produce an array of yeast strains each expressing Rpn7-tDimer2 and a unique combination of tagged E2 and E3 proteins, as described²⁰. The resulting strains were cultivated overnight at 20 °C in YPD medium and diluted in low fluorescence medium³⁷ 3–4 h before imaging. Imaging was performed in 8-well LabTek chambers or 96-well plates (Imaging plates CG, Zell-Kontakt) using an inverted Leica SP8 confocal microscope. Images of the BiFC signal were collected using a 514 nm laser and a narrow band-pass filter (525–538 nm) around the emission peak of the Venus fluorescent protein to reduce the contribution of cellular autofluorescence. Rpn7-tDimer2 was imaged simultaneously using a 580–630 nm filter. Cellular autofluorescence was imaged separately using the same band-pass filter as for BiFC images, but with a 458 nm excitation. Rpn7 localizes to the nucleus throughout the cell cycle in growing cells and relocates to cytoplasmic structures when cells enter quiescence³⁸. Rpn7-tDimer2 images were visually inspected before image processing to verify that cells are not quiescent. Rpn7-tDimer2 and autofluorescence images were used to segment the BiFC images into nuclear and cytoplasmic (whole cell minus nucleus) regions and to unmix the BiFC signal. Image segmentation and single-cell fluorescence measurements were performed using custom plugins in ImageJ³⁹ (available on request). To enable comparison of data from different experiments, the quantification results were rescaled so that BiFC signals of control cells had a mean of zero and a standard deviation of one. Statistical analysis and graphical representation were performed with GraphPad Prism software. Statistically significant differences from control cells were identified by one-way ANOVA followed by Bonferroni post-hoc tests to correct for multiple comparisons. No statistical method was used to predetermine sample size.

Recombinant protein expression and purification. *Escherichia coli* BL21(DE3) were transformed with plasmids encoding MBP-Hrd1^{CT} (Hrd1 residues 321–551), MBP-Asi1^{RING} (Asi1 residues 559–624), MBP-Asi3^{RING} (Asi3 residues 613–676), glutathione S-transferase (GST)-Ubc6^{ATM} (Ubc6 residues 1–230), GST-Ubc7 or Cue1^{U7BR} (Cue1 residues 151–203) and were cultivated in LB medium. Cue1^{U7BR} was coexpressed with GST-Ubc7. Protein expression was induced by addition of 1 mM isopropyl- β -D-thiogalactoside (IPTG) during 4 h at 25 °C. Cells were pelleted, resuspended in PBS, and lysed by sonication. Lysates were rotated with glutathione (GE Healthcare) or amylose beads (New England Biolab) for 1 h at 4 °C. Beads were washed with PBS containing 1 mM dithiothreitol (DTT). E2s were cleaved from GST using thrombin (Stago). MBP-E3s were eluted using 10 mM amylose and dialysed against PBS plus 1 mM DTT. All recombinant proteins were concentrated using spin filters (3 kDa, Amicon). Protein purity was tested by Coomassie staining after SDS-PAGE. Protein concentration was estimated by absorbance at 280 nm.

Microscale thermophoresis. Microscale thermophoresis analysis was performed essentially as described¹² using MBP-Asi1^{RING}, MBP-Asi3^{RING} or MBP-Hrd1^{CT} fluorescently labelled with the fluorescent dye NT-647 (labelling was performed with the Monolith Protein Labelling Kit RED-NHS according to the instructions of the supplier) and high precision standard treated capillaries. MBP-E3s were diluted to 100 nM in PBS, 5% glycerol, 0.1% Tween 20, 1 mM DTT, 10 μ M ZnAc and titrated with varying concentrations of unlabelled E2s before loading into capillaries. The difference of the thermophoretic properties of MBP-E3s were measured using a Monolith NT.115 instrument (NanoTemper Technologies GmbH) and a laser power of 60%. A nonlinear fit with the law of mass action was used to derive the dissociation constant (K_d) of the interaction as well as the theoretical thermophoretic properties of the MBP-E3 in its fully bound and unbound states. Those values were then used to normalize the measurements and calculate the

fraction of E3 bound at each E2 concentration. Data were plotted and fitted with the GraphPad Prism software.

tFT library construction. A total of 4,081 verified or uncharacterized *S. cerevisiae* open reading frames were selected for tagging based on structural and functional criteria (detailed in Supplementary Methods) to increase the probability that the C-terminal tFT tag would not affect protein functionality, and to avoid exposing the tag to an environment that could affect folding and maturation of the fluorescent proteins. Protocols for strain construction and validation are described in the Supplementary Methods. In brief, strain manipulations were automated and performed in 96-well format whenever possible. Using PCR targeting³⁰ and lithium acetate transformation of yeast⁴⁰, the module for seamless protein tagging with the mCherry-sfGFP timer (pMaM168 in Supplementary Table 5) was integrated into each selected genomic locus in the strain yMaM330 (Supplementary Table 4), a strain compatible with automated yeast genetics that carried a construct for conditional expression of the I-SceI meganuclease from the *GAL1* promoter integrated into the *leu2* locus. Correct integration of the tagging module into each locus and expression of tFT protein fusions was verified by PCR and whole colony fluorescence measurements for 4,044 open reading frames, with two independent clones validated for 3,952 open reading frames (Supplementary Table 2).

tFT library screening. Haploid array strains carrying deletions of individual components of the ubiquitin–proteasome system were obtained from the genome-wide heterozygous diploid yeast deletion library⁴¹ by sporulation and tetrad dissection. Screens were conducted in 1536-colony format. Using pinning robots (BioMatrix, S&P Robotics), tFT query strains (before marker excision) were mated with array mutants. Selection of diploids, sporulation and selection of haploids carrying simultaneously a tFT protein fusion and a gene deletion were performed by sequential pinning on appropriate selective media, as described²⁰, followed by seamless marker excision¹⁹. In each screen, a single tFT strain was crossed to a set of mutants in the ubiquitin–proteasome pathway (including the *asi1Δ*, *asi3Δ*, *hrd1Δ*, *doa10Δ*, *ubc6Δ* and *ubc7Δ* mutants) (A.K. *et al.*, manuscript in preparation) with four technical replicates of each cross. Technical replicates were arranged next to each other. Fluorescence intensities of the final colonies were measured after 24 h of growth on synthetic complete medium lacking histidine at 30 °C using Infinite M1000 or Infinite M1000 Pro plate readers equipped with stackers for automated plate loading (Tecan) and custom temperature control chambers. Measurements in mCherry (587/10 nm excitation, 610/10 nm emission, optimal detector gain) and sfGFP (488/10 nm excitation, 510/10 nm emission, optimal detector gain) channels were performed at 400 Hz frequency of the flash lamp, with ten flashes averaged for each measurement.

Measurements were filtered for potentially failed crosses based on colony size after haploid selection. Fluorescence intensity measurements were log-transformed and the data were normalized for spatial effects on plates by local regression. To estimate the changes from normal protein stability, median effects for tFT and deletion strains were subtracted from log-ratios of mCherry and sfGFP intensities. To avoid variance–mean dependences, standard deviations were regressed against the absolute fluorescence intensities. Changes in protein stability were divided by the regressed standard deviations, yielding a measurement comparable to a *z*-score, and tested against the hypothesis of zero change. A moderated *t*-test implemented in the R/Bioconductor package limma⁴² was used to compute *P* values. *P* values were adjusted for multiple testing by controlling the false discovery rate using the method of Benjamini–Hochberg.

Crosses with additional mutants were performed with independently constructed deletion strains using identical procedures on a RoToR pinning robot (Singer). Whole colony fluorescence intensities were corrected for autofluorescence using measurements of corresponding mutant colonies crossed to strain yMaM344-2 (Supplementary Table 4) expressing a truncated non-fluorescent mCherry^{ΔN} protein. For each tFT fusion, mCherry/sfGFP intensity ratios in each mutant were compared to a control cross with a wild-type strain carrying the *kanMX* selection marker in the *his3Δ* locus.

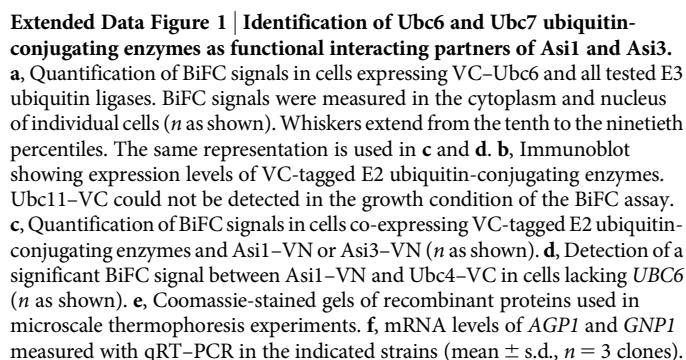
Fluorescence microscopy. Strains were grown at 30 °C in low fluorescence medium (synthetic complete medium prepared with yeast nitrogen base lacking folic acid and riboflavin; CYN6501, ForMedium) to $0.4\text{--}1.2 \times 10^7$ cells ml^{−1} and attached

to glass-bottom 96-well plates (MGB096-1-2-LG-L, Matrical) using Concanavalin A (C7275, Sigma) as described⁴³. Single plane images were acquired on a DeltaVision Elite system (Applied Precision) consisting of an inverted epifluorescence microscope (IX71; Olympus) equipped with an LED light engine (SpectraX, Lumencor), 475/28 and 575/25 nm excitation, and 525/50 and 624/40 nm emission filters (Semrock), a dual-band beam splitter 89021 (Chroma Technology), using either a 100× numerical aperture (NA) 1.4 UPlanSApo or a 60× NA 1.42 PlanApoN oil immersion objective (Olympus), an sCMOS camera (pco.edge 4.2, PCO) and a motorized stage contained in a temperature-controlled chamber. Image correction and quantification were performed in ImageJ³⁹. Dark signal and flat field corrections were applied to all images as described⁴³. Image deconvolution was performed with Softworx software (Applied Precision) using the conservative ratio algorithm with default parameter settings. Individual cell, perinuclear region and cytoplasm segmentation masks were manually defined in deconvolved images and applied to non-deconvolved images. Mean single-cell fluorescence measurements were corrected for cellular autofluorescence. Mean perinuclear fluorescence measurements were corrected for cytoplasmic fluorescence of each individual cell.

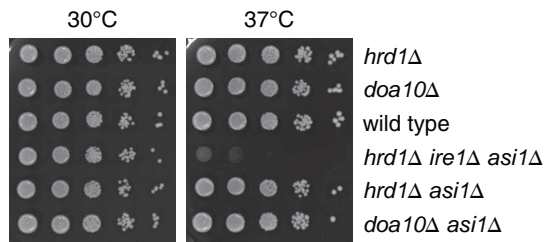
Strains expressing N- and C-terminally tagged Vtc1 and Vtc4 were imaged with exposure setting adjusted to the expression levels: 3.3-fold longer exposure time for C-terminally tagged fusions. Representative deconvolved images were scaled identically.

Cycloheximide chases. Strains were grown at 30 °C in synthetic complete medium to $\sim 0.8 \times 10^7$ cells ml^{−1} density before addition of cycloheximide to 100 μg ml^{−1} final concentration. One-millilitre samples taken at each time point were immediately mixed with 150 μl of 1.85 M NaOH and 10 μl β-mercaptoethanol, and flash frozen in liquid nitrogen. Whole-cell extracts were prepared as previously described⁴⁰, separated by SDS–PAGE followed by semi-dry blotting and probed sequentially with mouse anti-HA (12CA5) and mouse anti-Pgk1 (22C5D8, Molecular Probes) antibodies. A secondary goat anti-mouse antibody (IgG (H+L)-HRP, Dianova) was used for detection on a LAS-4000 system (Fuji).

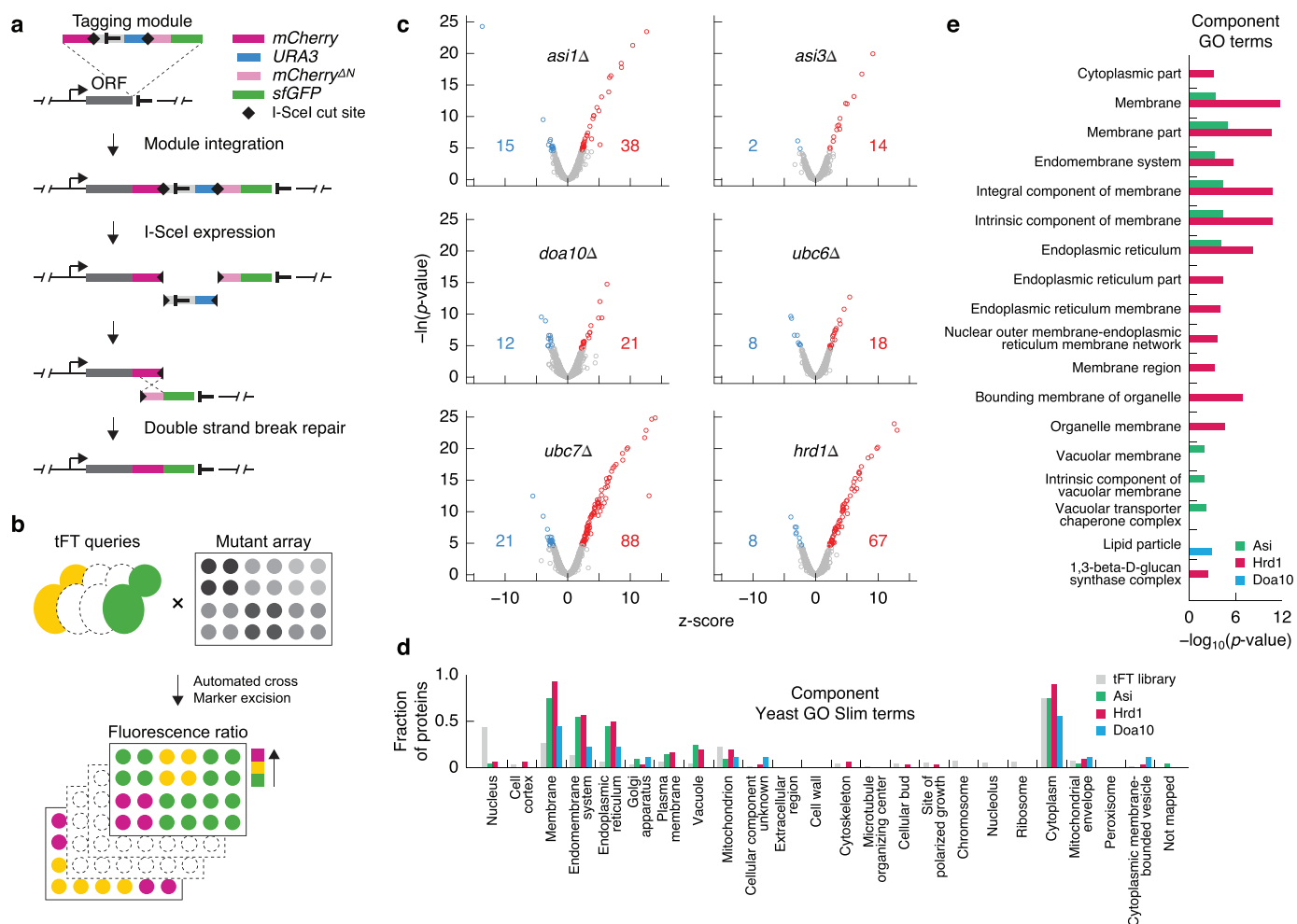
30. Janke, C. *et al.* A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**, 947–962 (2004).
31. Andréasson, C. & Ljungdahl, P. O. The N-terminal regulatory domain of Stp1p is modular and, fused to an artificial transcription factor, confers full Ssy1p–Ptr3p–Ssy5p sensor control. *Mol. Cell. Biol.* **24**, 7503–7513 (2004).
32. Becuwe, M. *et al.* A molecular switch on an arrestin-like protein relays glucose signaling to transporter endocytosis. *J. Cell Biol.* **196**, 247–259 (2012).
33. Léon, S., Erpapazoglou, Z. & Haguénauer-Tsapis, R. Ear1p and Ssh4p are new adaptors of the ubiquitin ligase Rsp5p for cargo ubiquitylation and sorting at multivesicular bodies. *Mol. Biol. Cell* **19**, 2379–2388 (2008).
34. Shyu, Y. J., Liu, H., Deng, X. & Hu, C. Identification of new fluorescent protein fragments for biomolecular fluorescence complementation analysis under physiological conditions. *Biotechniques* **40**, 61 (2006).
35. Sommer, T. & Jentsch, S. A protein translocation defect linked to ubiquitin conjugation at the endoplasmic reticulum. *Nature* **365**, 176–179 (1993).
36. Metzger, M. B. *et al.* A structurally unique E2-binding domain activates ubiquitination by the ERAD E2, Ubc7p, through multiple mechanisms. *Mol. Cell* **50**, 516–527 (2013).
37. Sheff, M. A. & Thorn, K. S. Optimized cassettes for fluorescent protein tagging in *Saccharomyces cerevisiae*. *Yeast* **21**, 661–670 (2004).
38. Laporte, D., Salin, B., Daignan-Fornier, B. & Sagot, I. Reversible cytoplasmic localization of the proteasome in quiescent yeast cells. *J. Cell Biol.* **181**, 737–745 (2008).
39. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675 (2012).
40. Knop, M. *et al.* Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. *Yeast* **15**, 963–972 (1999).
41. Winzler, E. A. *et al.* Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906 (1999).
42. Smyth, G. K. in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* (eds Gentleman, R., Carey, V., Dudoit, S., Irizarry, R. & Huber, W.) 397–420 (Springer, 2005).
43. Khmelinskii, A. & Knop, M. Analysis of protein dynamics with tandem fluorescent protein timers. *Methods Mol. Biol.* **1174**, 195–210 (2014).



©2014 Macmillan Publishers Limited. All rights reserved

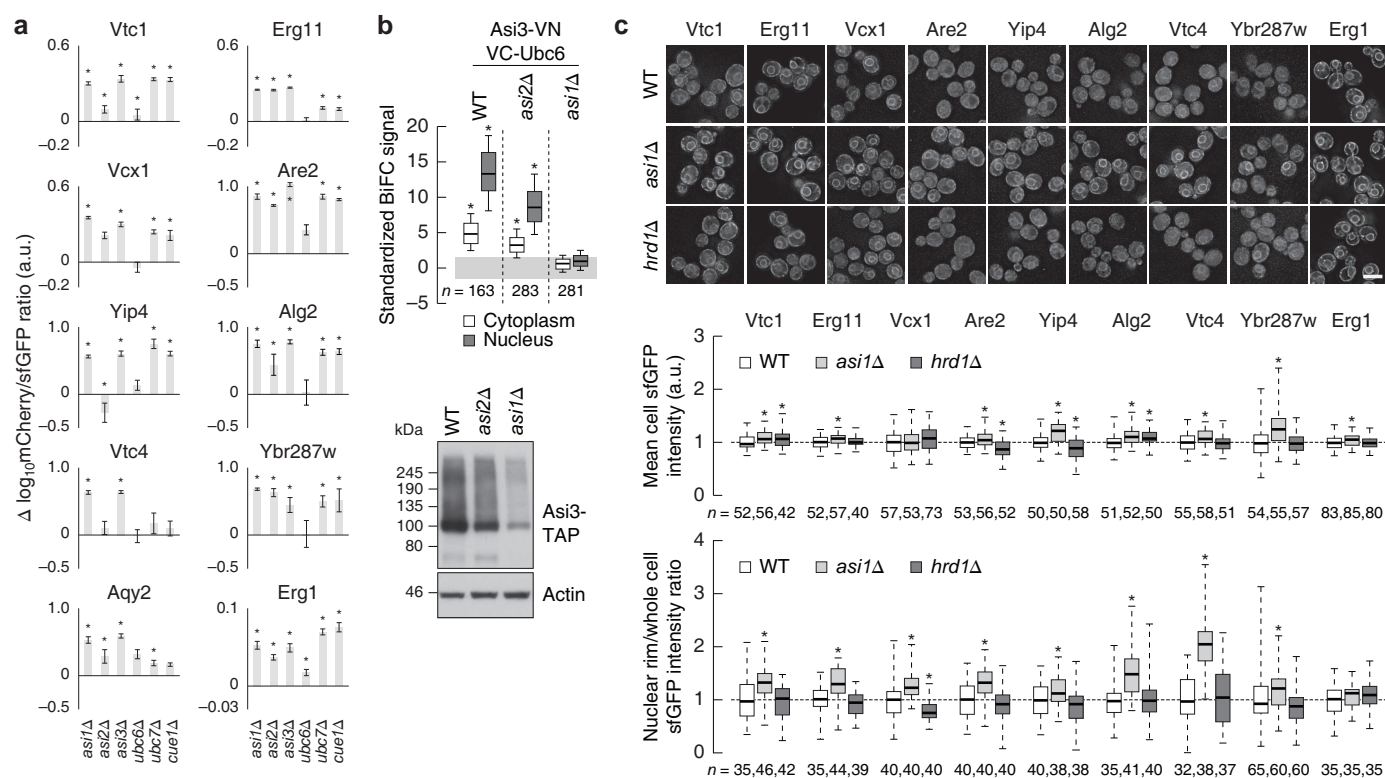


Extended Data Figure 2 | Lack of genetic interaction between *ASII* and *HRD1* or *DOA10* at 37 °C. Tenfold serial dilutions of strains grown on synthetic complete medium for 2 days at 30 or 37 °C.



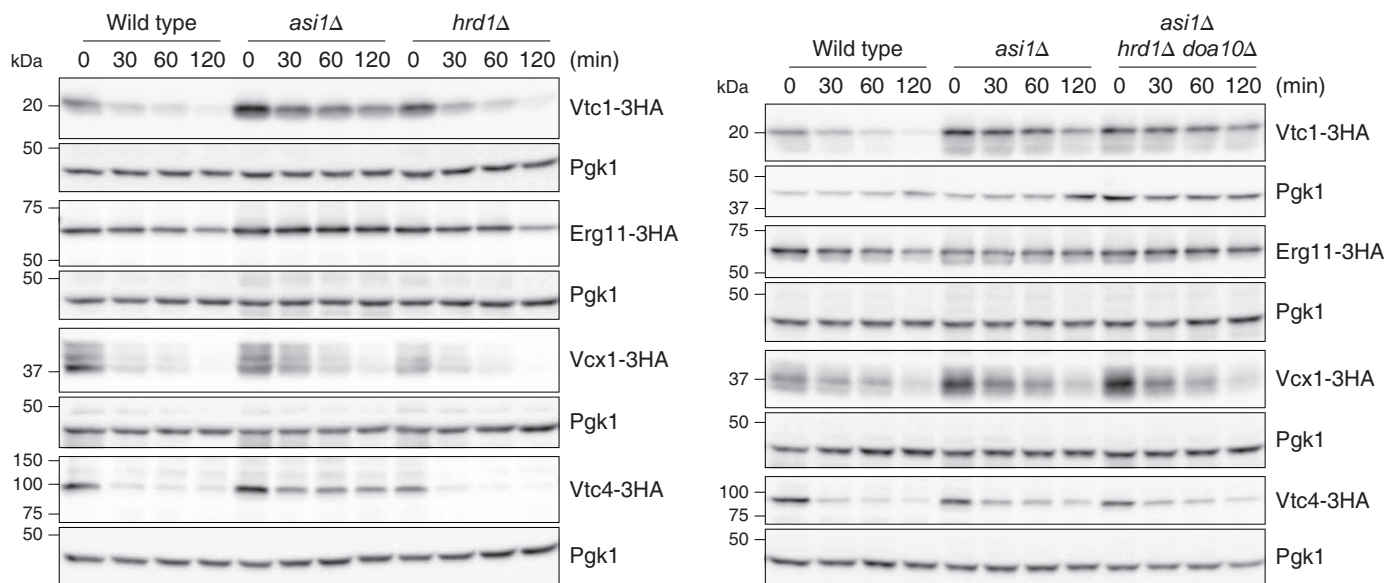
Extended Data Figure 3 | tFT screens for substrates of Asi and ERAD E3 ubiquitin ligases. **a**, Tagging approach used to construct the tFT library in a strain carrying the I-SceI meganuclease under an inducible promoter. First, a module for seamless C-terminal protein tagging with the mCherry-sfGFP timer is integrated into a genomic locus of interest using conventional PCR targeting. Subsequent I-SceI expression leads to excision of the heterologous terminator and the *URA3* selection marker, followed by repair of the double-strand break by homologous recombination between the mCherry and mCherry^{ΔN} sequences. A tFT fusion protein is expressed under control of endogenous promoter and terminator in the final strain. **b**, Workflow of screens for substrates of E3 ubiquitin ligases involved in protein degradation. Each tFT query strain is crossed to an array of mutants carrying different gene

deletion alleles. The resulting strains are imaged with a fluorescence plate reader to identify proteins with altered stability in each mutant. **c**, Volcano plots of the screens for proteins with altered stability in the indicated mutants. Plots show z-scores for changes in protein stability on the x axis and the negative logarithm of *P* values adjusted for multiple testing on the y axis. The number of proteins with increased (red) or decreased (blue) stability at 1% false discovery rate is indicated. **d**, Fraction of proteins in the tFT library and in the three clusters in Fig. 3b mapped to the full yeast slim set of component GO terms. Note that the GO term cytoplasm contains all cellular contents except the nucleus and the plasma membrane. **e**, The three clusters in Fig. 3b are enriched for proteins in the indicated component GO terms. Bar plot shows $-\log_{10}$ -transformed *P* values of significant enrichments.



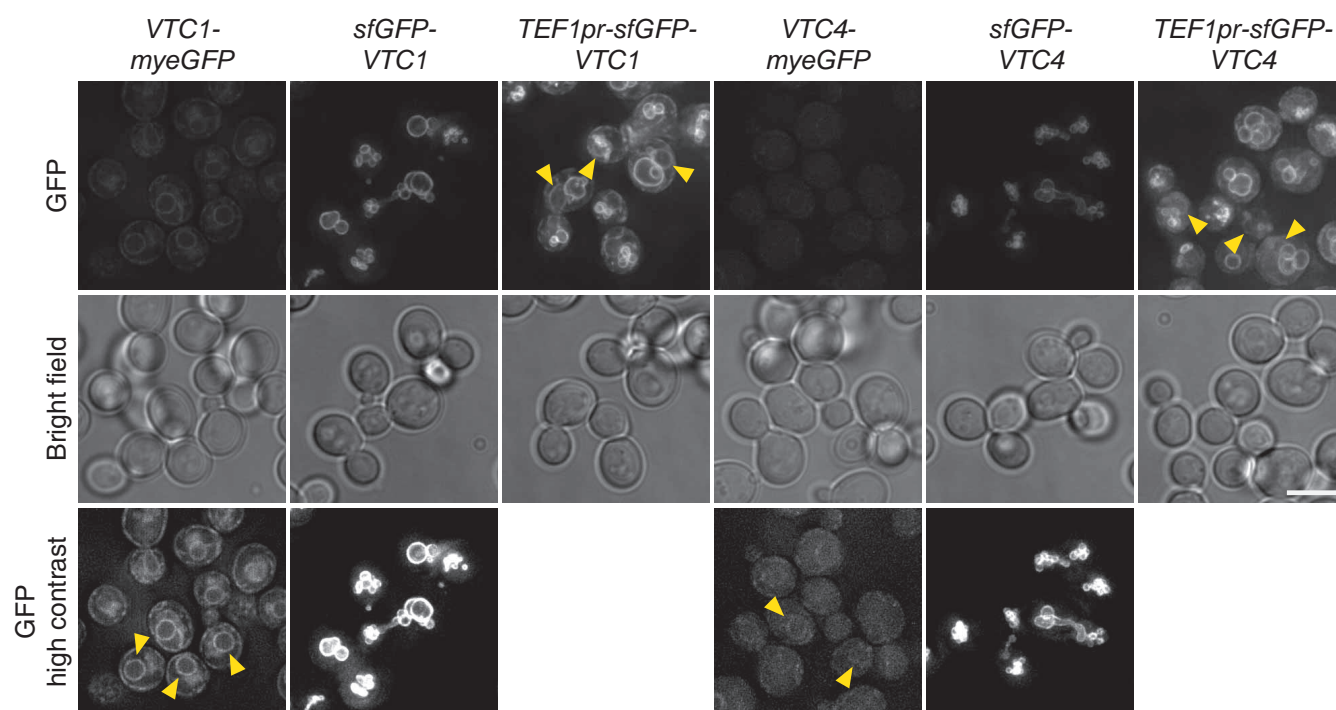
Extended Data Figure 4 | Analysis of integral membrane protein substrates of the Asi E3 ubiquitin ligase. **a**, Differences in the \log_{10} mCherry/sfGFP intensity ratio between the indicated mutants and the wild type (mean \pm s.d., $n = 4$) for tFT-tagged proteins from the Asi cluster in Fig. 3b. **b**, Quantification of BiFC signals in strains co-expressing VC-Ubc6 and Asi3-VN (top). BiFC signals were measured in the cytoplasm and nucleus of individual cells (n as shown). Whiskers extend from tenth to ninetieth percentiles. A substantial BiFC signal is retained in the *asi2Δ* mutant, despite reduced expression of Asi3

(immunoblot, bottom). **c**, Quantification of sfGFP signals in strains expressing tFT-tagged proteins from the Asi cluster in Fig. 3b. Fluorescence microscopy examples representative of five fields of view (top). Scale bar, 5 μm . sfGFP intensities were measured in individual cells (middle) and at the nuclear rim (bottom). For each protein, measurements were normalized to the mean of the respective wild type. Whiskers extend from minimum to maximum values. $*P < 0.05$ (a and c; two-tailed t -test) and $*P < 10^{-4}$ (b; one-way ANOVA with Bonferroni correction for multiple testing).



Extended Data Figure 5 | Cycloheximide chase experiments with substrates of the Asi E3 ubiquitin ligase. Degradation of 3×HA-tagged proteins after blocking translation with cycloheximide. Whole-cell extracts were separated by

SDS-PAGE followed by immunoblotting with antibodies against the HA tag and Pgk1 as loading control. Representative immunoblots from two technical replicates. Left, wild-type and *asi1Δ* immunoblots are reproduced in Fig. 3f.



Extended Data Figure 6 | Influence of tagging and expression levels on localization of Vtc1 and Vtc4. Fluorescence microscopy of strains expressing Vtc1 or Vtc4 tagged endogenously with monomeric yeast codon-optimized enhanced GFP (myeGFP) at the C terminus or tagged with sfGFP at the N

terminus and expressed under control of endogenous or *TEF1* promoters. Representative deconvolved images of five fields of view with ~100 cells each. Arrowheads indicate nuclear rim localization. Scale bar, 5 μ m.

Mitochondrial UPR-regulated innate immunity provides resistance to pathogen infection

Mark W. Pellegrino¹, Amrita M. Nargund¹, Natalia V. Kirienko^{2,3}, Reba Gillis¹, Christopher J. Fiorese⁴ & Cole M. Haynes^{1,4}

Metazoans identify and eliminate bacterial pathogens in microbe-rich environments such as the intestinal lumen; however, the mechanisms are unclear. Host cells could potentially use intracellular surveillance or stress response programs to detect pathogens that target monitored cellular activities and then initiate innate immune responses^{1–3}. Mitochondrial function is evaluated by monitoring mitochondrial protein import efficiency of the transcription factor ATFS-1, which mediates the mitochondrial unfolded protein response (UPR^{mt}). During mitochondrial stress, mitochondrial import is impaired⁴, allowing ATFS-1 to traffic to the nucleus where it mediates a transcriptional response to re-establish mitochondrial homeostasis⁵. Here we examined the role of ATFS-1 in *Caenorhabditis elegans* during pathogen exposure, because during mitochondrial stress ATFS-1 induced not only mitochondrial protective genes but also innate immune genes that included a secreted lysozyme and anti-microbial peptides. Exposure to the pathogen *Pseudomonas aeruginosa* caused mitochondrial dysfunction and activation of the UPR^{mt}. *C. elegans* lacking *atfs-1* were susceptible to *P. aeruginosa*, whereas hyper-activation of ATFS-1 and the UPR^{mt} improved clearance of *P. aeruginosa* from the intestine and prolonged *C. elegans* survival in a manner mainly independent of known innate immune pathways^{6,7}. We propose that ATFS-1 import efficiency and the UPR^{mt} is a means to detect pathogens that target mitochondria and initiate a protective innate immune response.

Animals harbour bacteria that are essential for normal physiology⁸; however, they must distinguish between commensal and pathogenic microbes to maintain homeostasis. Pathogenic bacteria can be recognized directly or by damage inflicted by the pathogen⁹ leading to activation of innate immunity responses that limit pathogen growth. Recently it has been demonstrated that perturbations to protein synthesis, proteolysis or mitochondrial activity are sufficient to activate innate immune responses, suggesting the elegant hypothesis that host cells use intracellular stress responses to initiate innate immunity programs when pathogens perturb monitored cellular processes^{1–3}.

Cells respond to mitochondrial dysfunction by activating the UPR^{mt}, which is regulated by the transcription factor ATFS-1. In healthy cells, ATFS-1 is efficiently imported into mitochondria and degraded. However, during mitochondrial stress, mitochondrial import efficiency is reduced^{4,5}, allowing a small percentage of ATFS-1 to accumulate in the cytosol⁵. Because ATFS-1 has a nuclear localization sequence (NLS), it then traffics to the nucleus where it activates a protective transcriptional response (Fig. 1a). Our expression profiling studies indicated that ATFS-1 induces genes that promote mitochondrial protein folding, reactive oxygen species (ROS) detoxification and mitochondrial protein import, suggesting the UPR^{mt} stabilizes the mitochondrial protein folding environment to promote organelle homeostasis⁵.

Intriguingly, a number of transcripts induced during mitochondrial stress caused by inhibition of the mitochondrial protease SPG-7 encode innate immunity proteins⁵ (Extended Data Table 1), some of which were also found to be induced following exposure to the pathogen *P. aeruginosa*¹⁰ (Fig. 1b and Extended Data Table 2). The antimicrobial peptide

abf-2 and the secreted lysozyme *lys-2*, both of which are required for resistance to pathogen infection^{11,12}, were induced during mitochondrial stress (Fig. 1c, d), as were two C-type lectins, which are involved in pathogen recognition¹³ (Fig. 1e, f). Mitochondrial-specific stress also caused induction of antimicrobial peptides¹⁴ in mammalian cells (Fig. 1g–j), suggesting the response is conserved. In *C. elegans*, induction of innate immune genes by *spg-7* RNA interference (*spg-7*(RNAi)) required ATFS-1 (Fig. 1c–f). Thus, in addition to inducing mitochondrial-protective genes, ATFS-1 also transcriptionally upregulated innate immune genes during mitochondrial stress. Therefore we hypothesized that ATFS-1 and the UPR^{mt} are involved in regulating innate immunity during exposure to pathogens that perturb mitochondrial function.

P. aeruginosa produces virulence factors that target many cellular functions including the mitochondrial toxins cyanide and pyocyanin^{15,16}. *P. aeruginosa* also produces exotoxin A, which impairs protein synthesis and leads to the induction of the innate immune gene *irg-1* via the transcription factor ZIP-2 (refs 2, 3, 17). Mitochondrial stress also caused *irg-1*_{pr::gfp} (pr, promoter) induction, which was blocked in *atfs-1(tm4919)* and partially so in *zip-2(tm4248)* worms (Fig. 1k), suggesting that multiple transcription factors and stressors influence innate immune gene expression. *zip-2* mRNA was also induced during mitochondrial stress, which also required *atfs-1* (Fig. 1l). *F35E12.5*, which is induced by the MAP kinase PMK-1 and the transcription factor ATF-7 during *P. aeruginosa* exposure^{10,18}, was not induced during mitochondrial stress (Extended Data Fig. 1a). Thus, ATFS-1 regulates a subset of innate immune genes during mitochondrial stress in addition to its cytoprotective role in promoting mitochondrial homeostasis.

We next examined if *P. aeruginosa* exposure caused mitochondrial stress capable of activating the UPR^{mt}. Slow-killing conditions were used in which the pathogen accumulates in the intestine leading to infection¹⁹. Interestingly, *P. aeruginosa* exposure caused intestinal cell mitochondria to elongate similar to *spg-7*(RNAi) treatment (Fig. 2a), consistent with the pathogen causing mitochondrial stress, and mitochondrial fusion providing protection²⁰. Exposure to *P. aeruginosa* also caused striking developmental delays in combination with mild mitochondrial stresses such as ethidium bromide⁵, paraquat⁵ or the *clk-1(qm30)* allele²¹ (Fig. 2b), consistent with the pathogen causing modest mitochondrial stress. Importantly, *P. aeruginosa* exposure caused an *atfs-1*-dependent increase in mitochondrial chaperone reporter (*hsp-6* and *hsp-60_{pr::gfp}*) activation in the intestine (Fig. 2c and Extended Data Fig. 1b), which correlated with increased nuclear accumulation of ATFS-1::GFP and required the NLS in ATFS-1 (Fig. 2d and Extended Data Fig. 1c, d). Exposure to *P. aeruginosa* liquid-killing conditions, which requires pathogen-expressed iron chelating siderophores²², also induced mitochondrial chaperone genes, suggesting multiple *P. aeruginosa* virulence factors can activate the UPR^{mt} (Extended Data Fig. 2a, b). Interestingly, both synthetic growth arrest and UPR^{mt} activation required the *P. aeruginosa* global virulence activator gene *gacA*²³ (Fig. 2b, c). Furthermore, exposure to *P. aeruginosa* strains lacking individual siderophore, pyocyanin or cyanide toxin genes resulted in less UPR^{mt} activation than wild-type *P. aeruginosa*

¹Cell Biology Program, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, New York 10065, USA. ²Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. ³Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁴BCMB Allied Program, Weill Cornell Medical College, 1300 York Avenue, New York, New York 10065, USA.

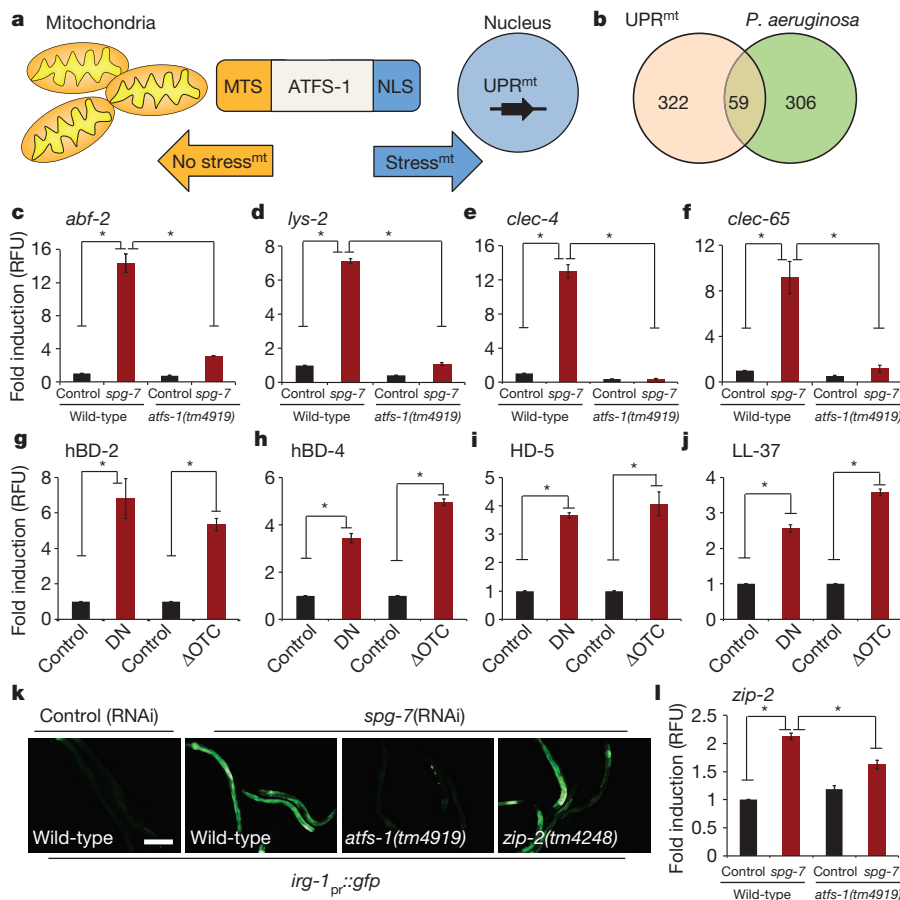


Figure 1 | ATFS-1 induces innate immunity genes during mitochondrial dysfunction.

a, Schematic of UPR^{mt} regulation. **b**, Diagram of ATFS-1-dependent UPR^{mt} genes⁵ in common with genes induced by *P. aeruginosa*¹⁰. **c–f**, *abf-2*, *lys-2*, *clec-4* and *clec-65* transcripts as determined by qRT-PCR in wild-type or *atfs-1(tm4919)* worms on control versus *spg-7(RNAi)* ($n = 3$, \pm s.d.), $*P < 0.05$ (Student's *t*-test). **g–j**, Antimicrobial peptide transcripts in human cells during mitochondrial stress caused by expression of dominant-negative AFG3L2 (DN) or misfolded ornithine transcarbamylase (Δ OTC) as determined by qRT-PCR ($n = 3$, \pm s.d.), $*P < 0.05$ (Student's *t*-test). **k**, *irg-1_{pr}::gfp* in wild-type, *atfs-1(tm4919)* or *zip-2(tm4248)* worms on control versus *spg-7(RNAi)*. Scale bar, 0.15 mm. **l**, *zip-2* transcripts as determined by qRT-PCR in wild-type or *atfs-1(tm4919)* worms on control or *spg-7(RNAi)* ($n = 3$, \pm s.d.), $*P < 0.05$ (Student's *t*-test). RFU, relative fluorescence units.

(Extended Data Fig. 2d, e), suggesting that multiple pathogen toxins target mitochondrial function resulting in UPR^{mt} activation. However, UPR^{mt} activation may also be due to indirect damage associated with activation of a separate immune response²⁴.

We examined the role of ATFS-1 in the induction of innate immune genes during *P. aeruginosa* exposure rather than specifically during mitochondrial stress. Similarly, *abf-2*, *lys-2*, *clec-4* and *clec-65* were induced upon *P. aeruginosa* exposure independent of exogenous mitochondrial stress, which also required *atfs-1* (Fig. 2e–h). Similar to the mitochondrial chaperones, both *lys-2_{pr}::gfp* and *irg-1_{pr}::gfp* were induced in the intestine upon *P. aeruginosa* exposure (Fig. 2i and Extended Data Fig. 3). Interestingly, increased *irg-1_{pr}::gfp* expression was impaired in both *atfs-1(tm4919)* and *zip-2(tm4248)* mutants (Fig. 2i). Furthermore, *zip-2* transcript induction on *P. aeruginosa*¹⁷ was also partially impaired in *atfs-1* mutant worms, suggesting *atfs-1* can function upstream of *zip-2* (Extended Data Fig. 4a).

Consistent with a role for ATFS-1 in inducing innate immune and mitochondrial protective genes⁵, the survival of worms raised on *atfs-1(RNAi)* was significantly reduced when exposed to *P. aeruginosa*, but not *E. coli* (Fig. 3a, b). *atfs-1(RNAi)* treated worms were also susceptible to *P. aeruginosa* liquid-killing (Extended Data Fig. 2c), supporting a role for ATFS-1 in activating a protective transcriptional response to pathogen exposure. RNAi was used to reduce *atfs-1* activity for the survival studies rather than *atfs-1(tm4919)* because of germline defects that complicate the analysis (Extended Data Fig. 4b, c).

We examined if UPR^{mt} activation is sufficient to protect against *P. aeruginosa*. The UPR^{mt} was induced by allowing worms to develop on *spg-7(RNAi)* for two days⁵ before pathogen exposure. UPR^{mt} pre-activation dramatically reduced the intestinal accumulation of *P. aeruginosa* expressing GFP (*P. aeruginosa*-GFP¹⁹) (Fig. 3c, d). Importantly, *P. aeruginosa*-GFP accumulated in the intestine of *atfs-1(tm4919)* worms following

spg-7(RNAi) treatment indicating that UPR^{mt} activation promotes pathogen clearance. In addition to adapting transcription, worms are also able to avoid *P. aeruginosa*, which was unaffected by *atfs-1(tm4919)* or pre-treatment with *spg-7(RNAi)* (Extended Data Fig. 5a–e). Consistent with increased pathogen clearance via anti-microbial gene induction, UPR^{mt} pre-activation prolonged the survival of animals challenged with *P. aeruginosa*, which required *atfs-1* (Fig. 3e) and was independent of germline defects or feeding behaviour (Extended Data Fig. 5f, g).

Because mitochondrial stress can activate multiple stress response pathways in addition to the UPR^{mt} (refs 25, 26), we examined an *atfs-1* gain-of-function mutant, which constitutively activates the UPR^{mt} independent of mitochondrial dysfunction. *atfs-1(et18)* worms express ATFS-1 with an amino acid substitution in the mitochondria targeting sequence that reduces mitochondrial import efficiency causing constitutive UPR^{mt} activation²⁷ and innate immune gene induction (Extended Data Fig. 6a–e). We observed that *atfs-1(et18)* worms accumulated less *P. aeruginosa*-GFP in the intestine (Fig. 3f, g) and survived longer than wild-type worms (Fig. 3h) indicating that UPR^{mt} activation is sufficient to provide resistance to *P. aeruginosa*. Importantly, *atfs-1(RNAi)* and *lys-2(RNAi)* reduced *atfs-1(et18)* worm survival (Fig. 3h and Extended Data Fig. 6f), suggesting that ATFS-1-mediated innate immune gene induction provides resistance to *P. aeruginosa*.

Inhibition of additional cellular activities including translation (*eft-2*, also known as *eef-2*), mRNA splicing (*T08A11.2*), calcium transport (*sca-1*) and the pentose phosphate pathway (*T25B9.9*) also induce innate immune gene expression^{1–3} but do not induce the UPR^{mt} (Extended Data Fig. 7a, b). Thus, we examined if other stress-activated innate immune responses are also protective against *P. aeruginosa*. Knockdown of *eft-2*, *T25B9.9*, *sca-1* or *T08A11.2* did not increase survival on *P. aeruginosa* (Extended Data Fig. 7c), however, *sca-1(RNAi)* and *T08A11.2(RNAi)* decreased lifespan on *E. coli*, indicating a reduction in general fitness (Extended Data

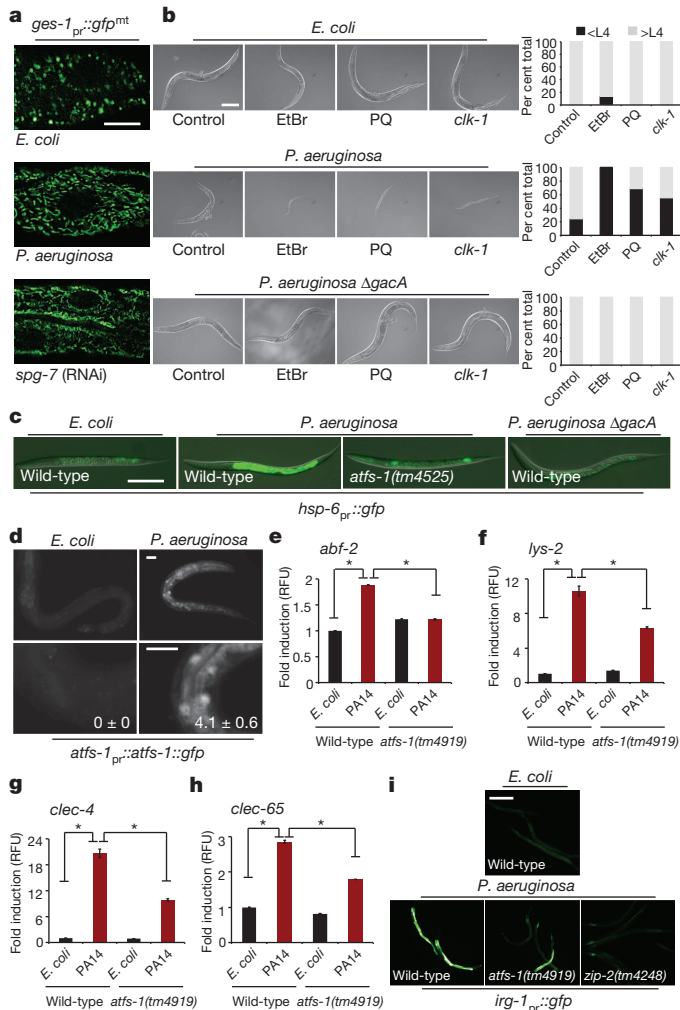


Figure 2 | Mitochondrial stress and UPR^{mt} activation by *P. aeruginosa*. **a**, *ges-1_{pr::gfp}^{mt}* intestinal cell mitochondria on *E. coli*, *P. aeruginosa* or *spg-7*(RNAi). Scale bar, 0.05 mm. **b**, Worms treated with ethidium bromide (EtBr), paraquat (PQ), and *clk-1*(*qm30*) worms raised on *E. coli*, *P. aeruginosa* or *P. aeruginosa* Δ *gacA*. Quantification of the developmental stage for each treatment shown next to the corresponding panel ($n = 35$ each treatment). Scale bar, 0.1 mm. **c**, Wild-type or *atfs-1*(*tm4919*);*hsp-6_{pr::gfp}* worms on *E. coli*, *P. aeruginosa* or *P. aeruginosa* Δ *gacA*. Scale bar, 0.1 mm. **d**, *atfs-1_{pr::atfs-1::gfp}* on *E. coli* or *P. aeruginosa*. Lower panels are higher magnification ($n = 3$). Mean percentages of ATFS-1::GFP nuclear accumulation are indicated (\pm s.e.m.). Scale bars, 0.1 mm. **e–h**, *abf-2*, *lys-2*, *clec-4* and *clec-65* transcripts as determined by qRT-PCR in wild-type or *atfs-1*(*tm4919*) worms on *E. coli* or *P. aeruginosa* ($n = 3$, \pm s.d.), * $P < 0.05$ (Student's *t*-test). RFU, relative fluorescence units. **i**, Wild-type, *atfs-1*(*tm4919*) and *zip-2*(*tm4248*) *irg-1_{pr::gfp}* worms on *E. coli* or *P. aeruginosa*. Scale bar, 0.05 mm.

Fig. 7d). In contrast, knockdown of the mitochondrial ATP synthase subunit *atp-2*, which activates mitochondrial protective and innate immune gene expression (Extended Data Fig. 7a, b), prolonged survival during *P. aeruginosa* exposure (Extended Data Fig. 7c). Our data suggest the UPR^{mt} provides protection from *P. aeruginosa* by coupling mitochondrial-protective and antimicrobial gene expression.

Lastly, we determined if ATFS-1 and the UPR^{mt} interacted with established *C. elegans* innate immune pathways, which include a MAP kinase pathway mediated by NSY-1/SEK-1/PMK-1 (refs 6, 7, 10), the MLK-1/MEK-1/KGB-1 c-Jun kinase pathway^{7,28}, as well as that mediated by ZIP-2 (ref. 17). Interestingly, pre-activation of the UPR^{mt} enhanced the survival of the *pmk-1* and *sek-1* mutants (Fig. 4a, b), as well as the *kbg-1* and *mlk-1* mutants (Fig. 4c, d). Of note, increased survival by *spg-7*(RNAi) was further enhanced in *kbg-1*(*km21*) worms, consistent with *kbg-1* being a negative regulator of the UPR^{mt} (ref. 29) (Extended Data Fig. 7e). In

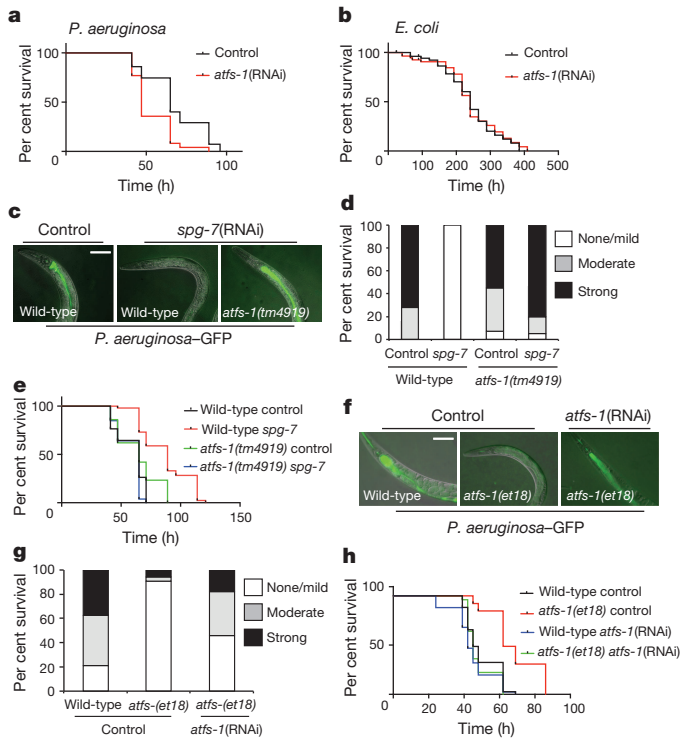


Figure 3 | UPR^{mt} activation provides resistance to *P. aeruginosa*. **a**, **b**, Survival of worms on control or *atfs-1*(RNAi) exposed to *P. aeruginosa* or *E. coli*. Statistics are in Extended Data Table 3. **c**, **d**, Images and quantification of *P. aeruginosa*-GFP in wild-type or *atfs-1*(*tm4919*) worms on control or *spg-7*(RNAi). Scale bar, 0.1 mm ($n = 35$ for each treatment). **e**, Survival of wild-type and *atfs-1*(*tm4919*) worms on control or *spg-7*(RNAi) exposed to *P. aeruginosa*. Statistics are in Extended Data Table 3. **f**, **g**, Images and quantification of *P. aeruginosa*-GFP in wild-type and *atfs-1*(*et18*) worms on control or *atfs-1*(RNAi) ($n = 35$ for each treatment). Scale bar, 0.1 mm. **h**, Survival of wild-type and *atfs-1*(*et18*) worms on control or *atfs-1*(RNAi) exposed to *P. aeruginosa*. Statistics are in Extended Data Table 3.

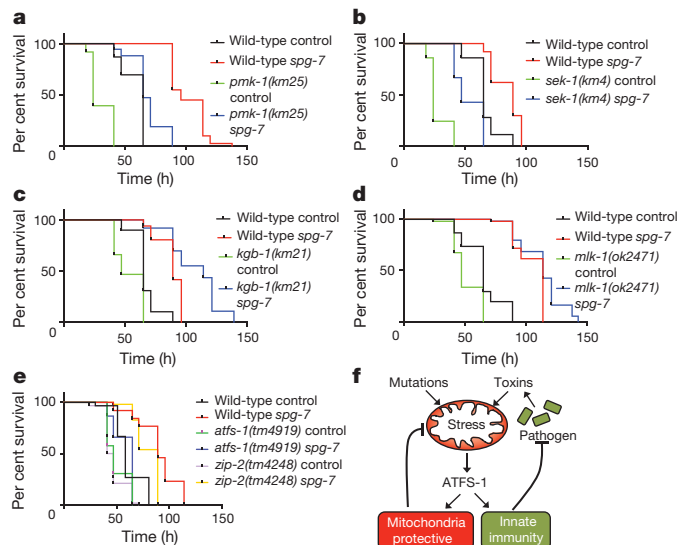


Figure 4 | UPR^{mt} activation prolongs survival independently of known innate immune pathways. **a–d**, Survival of wild-type, *pmk-1*(*km25*), *sek-1*(*km4*), *kbg-1*(*km21*) and *mlk-1*(*ok2471*) worms on control or *spg-7*(RNAi) exposed to *P. aeruginosa*. Statistics are in Extended Data Table 3. **e**, Survival of wild-type, *zip-2*(*tm4248*) and *atfs-1*(*tm4919*) worms raised on control or *spg-7*(RNAi) and exposed to *P. aeruginosa*. Statistics are in Extended Data Table 3. **f**, ATFS-1 signalling schematic.

contrast, *zip-2(tm4248)* modestly reduced the enhanced resistance conferred by *spg-7(RNAi)* (Fig. 4e), consistent with *atfs-1* functioning in the same pathway as *zip-2* during mitochondrial stress. Together, our data suggest that the UPR^{mt} can function independently of the MAP and c-Jun kinase regulated innate immune pathways.

Our studies indicate that the UPR^{mt} is activated by and protects against *P. aeruginosa*, and thus support a mechanistic means⁵ by which host cells can detect pathogens that target mitochondrial function (Fig. 4f), which is consistent with only a subset of bacterial species inducing the UPR^{mt} (ref. 30). Because ATFS-1 responds directly to mitochondrial dysfunction and induces a transcriptional response that is both mitochondrial protective⁵ and antimicrobial, the UPR^{mt} is a uniquely positioned pathway to mitigate mitochondrial damage stemming from genetic defects or pathogen exposure (Fig. 4f).

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 March; accepted 14 August 2014.

Published online 28 September 2014.

- Melo, J. A. & Ruvkun, G. Inactivation of conserved *C. elegans* genes engages pathogen- and xenobiotic-associated defenses. *Cell* **149**, 452–466 (2012).
- McEwan, D. L., Kirienko, N. V. & Ausubel, F. M. Host translational inhibition by *Pseudomonas aeruginosa* Exotoxin A triggers an immune response in *Caenorhabditis elegans*. *Cell Host Microbe* **11**, 364–374 (2012).
- Dunbar, T. L., Yan, Z., Balla, K. M., Smelkinson, M. G. & Troemel, E. R. *C. elegans* detects pathogen-induced translational inhibition to activate immune signaling. *Cell Host Microbe* **11**, 375–386 (2012).
- Wright, G., Terada, K., Yano, M., Sergeev, I. & Mori, M. Oxidative stress inhibits the mitochondrial import of preproteins and leads to their degradation. *Exp. Cell Res.* **263**, 107–117 (2001).
- Nargund, A. M., Pellegrino, M. W., Fiorese, C. J., Baker, B. M. & Haynes, C. M. Mitochondrial import efficiency of ATFS-1 regulates mitochondrial UPR activation. *Science* **337**, 587–590 (2012).
- Kim, D. H. *et al.* A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. *Science* **297**, 623–626 (2002).
- Twumasi-Boateng, K. *et al.* An age-dependent reversal in the protective capacities of JNK signaling shortens *Caenorhabditis elegans* lifespan. *Aging Cell* **11**, 659–667 (2012).
- Guarner, F. & Malagelada, J. R. Gut flora in health and disease. *Lancet* **361**, 512–519 (2003).
- Newton, K. & Dixit, V. M. Signaling in innate immunity and inflammation. *Cold Spring Harb. Perspect. Biol.* **4**, <http://dx.doi.org/10.1101/cshperspecta.006049> (2012).
- Troemel, E. R. *et al.* p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans*. *PLoS Genet.* **2**, e183 (2006).
- Kato, Y. *et al.* *abf-1* and *abf-2*, ASABF-type antimicrobial peptide genes in *Caenorhabditis elegans*. *Biochem. J.* **361**, 221–230 (2002).
- Nandakumar, M. & Tan, M. W. Gamma-linolenic and stearidonic acids are required for basal immunity in *Caenorhabditis elegans* through their effects on p38 MAP kinase activity. *PLoS Genet.* **4**, e1000273 (2008).
- Nicholas, H. R. & Hodgkin, J. Responses to infection and possible recognition strategies in the innate immune system of *Caenorhabditis elegans*. *Mol. Immunol.* **41**, 479–493 (2004).
- De Smet, K. & Contreras, R. Human antimicrobial peptides: defensins, cathelicidins and histatins. *Biotechnol. Lett.* **27**, 1337–1347 (2005).
- O'Malley, Y. Q. *et al.* Subcellular localization of *Pseudomonas* pyocyanin cytotoxicity in human lung epithelial cells. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **284**, L420–L430 (2003).
- Gallagher, L. A. & Manoil, C. *Pseudomonas aeruginosa* PAO1 kills *Caenorhabditis elegans* by cyanide poisoning. *J. Bacteriol.* **183**, 6207–6214 (2001).
- Estes, K. A., Dunbar, T. L., Powell, J. R., Ausubel, F. M. & Troemel, E. R. bZIP transcription factor *zip-2* mediates an early response to *Pseudomonas aeruginosa* infection in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **107**, 2153–2158 (2010).
- Pukkila-Worley, R. *et al.* Stimulation of host immune defenses by a small molecule protects *C. elegans* from bacterial infection. *PLoS Genet.* **8**, e1002733 (2012).
- Tan, M. W., Mahajan-Miklos, S. & Ausubel, F. M. Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. *Proc. Natl Acad. Sci. USA* **96**, 715–720 (1999).
- Gomes, L. C., Di Benedetto, G. & Scorrano, L. During autophagy mitochondria elongate, are spared from degradation and sustain cell viability. *Nature Cell Biol.* **13**, 589–598 (2011).
- Feng, J., Bussiere, F. & Hekimi, S. Mitochondrial electron transport is a key determinant of life span in *Caenorhabditis elegans*. *Dev. Cell* **1**, 633–644 (2001).
- Kirienko, N. V. *et al.* *Pseudomonas aeruginosa* disrupts *Caenorhabditis elegans* iron homeostasis, causing a hypoxic response and death. *Cell Host Microbe* **13**, 406–416 (2013).
- Tan, M. W., Rahme, L. G., Sternberg, J. A., Tompkins, R. G. & Ausubel, F. M. *Pseudomonas aeruginosa* killing of *Caenorhabditis elegans* used to identify *P. aeruginosa* virulence factors. *Proc. Natl Acad. Sci. USA* **96**, 2408–2413 (1999).
- Richardson, C. E., Kooistra, T. & Kim, D. H. An essential role for XBP-1 in host protection against immune activation in *C. elegans*. *Nature* **463**, 1092–1095 (2010).
- Baker, B. M., Nargund, A. M., Sun, T. & Haynes, C. M. Protective coupling of mitochondrial function and protein synthesis via the eIF2 α kinase GCN-2. *PLoS Genet.* **8**, e1002760 (2012).
- Walter, L., Baruah, A., Chang, H. W., Pace, H. M. & Lee, S. S. The homeobox protein CEH-23 mediates prolonged longevity in response to impaired mitochondrial electron transport chain in *C. elegans*. *PLoS Biol.* **9**, e1001084 (2011).
- Rauthan, M., Ranji, P., Aguilera Pradenas, N., Pitot, C. & Pilon, M. The mitochondrial unfolded protein response activator ATFS-1 protects cells from inhibition of the mevalonate pathway. *Proc. Natl Acad. Sci. USA* **110**, 5981–5986 (2013).
- Kim, D. H. *et al.* Integration of *Caenorhabditis elegans* MAPK pathways mediating immunity and stress resistance by MEK-1 MAPK kinase and VHP-1 MAPK phosphatase. *Proc. Natl Acad. Sci. USA* **101**, 10990–10994 (2004).
- Runkel, E. D., Liu, S., Baumeister, R. & Schulze, E. Surveillance-activated defenses block the ROS-induced mitochondrial unfolded protein response. *PLoS Genet.* **9**, e1003346 (2013).
- Liu, Y., Samuel, B. S., Breen, P. C. & Ruvkun, G. *Caenorhabditis elegans* pathways that surveil and defend mitochondria. *Nature* **508**, 406–410 (2014).

Acknowledgements We thank E. Troemel, M. Pilon, S. Mitani of the National Bioresource Project, and the *Caenorhabditis* Genetics Center for providing *C. elegans* strains. We thank D. Kim and J. Xavier for providing *P. aeruginosa* strains, and T. Langer and S. Troeder for mammalian reagents. This work was supported by the Ellison Medical Foundation, the Lucille Castori Center for Microbes, Inflammation and Cancer at MSKCC and the National Institutes of Health (R01AG040061) to C.M.H., (F32AI100501) to N.V.K. and (R01AI085581) to F. Ausubel.

Author Contributions M.W.P., A.M.N., N.V.K., R.G. and C.J.F. performed experiments. M.W.P. and C.M.H. conceived of and planned the experiments and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.M.H. (haynec@mskcc.org).

METHODS

Worm and bacterial strains. The *atfs-1(tm4919)* mutant strain was a gift from the National BioResource Project and backcrossed twice to wild-type N2 worms. Worm strains were provided by the Caenorhabditis Genetics Center unless otherwise noted. Hermaphrodite worms were raised on the OP50 strain of *E. coli* unless they were treated with RNAi, in which case the HT115 *E. coli* strain expressing the described RNAi plasmid was used^{31,32}. Where indicated, worms were exposed to the pathogenic strain of *P. aeruginosa*, PA14.

Cell culture. Expression of dominant-negative AFG3L2 was induced in stable HEK293 cells by the addition of 1 $\mu\text{g ml}^{-1}$ tetracycline³³ and the cells were collected 48 h later. The ΔOTC expression plasmid³⁴ was transfected into HeLa cells via Lipofectamine and the cells were collected after 72 h.

C. elegans slow-killing assay. Slow-killing experiments were performed as previously described^{35,36} with minor modifications. *E. coli* or *P. aeruginosa* overnight cultures were used to seed slow-killing nematode growth medium (NGM) agar plates (with 0.35% peptone). Plates were allowed to dry overnight at room temperature, incubated at 37°C for 24 h and allowed to equilibrate at room temperature. Synchronized L1 worms were allowed to develop on *E. coli* until the L4 stage and then transferred to *P. aeruginosa* slow-killing plates and incubated at 25°C. RNAi was performed as described previously³⁷. For *atfs-1(RNAi)* (Fig. 3a, b), *eri-1(mg366)* (enhanced RNAi) worms³⁸ were raised on control or *atfs-1(RNAi)* bacteria at 16°C until the L4 stage. All animals were transferred to fresh *P. aeruginosa* slow-killing plates in a randomized fashion. Animals were counted at the described times and were scored as dead if they failed to respond when touched. Fifty worms were used per experiment and those that had crawled off the plate or exploded at the vulva were excluded. All data related to the survival analysis is presented in Extended Data Table 3. Each experiment was performed in triplicate and the log rank (Mantel–Cox) statistical test was used to evaluate *P* values.

Intestinal mitochondrial morphology was visualized using *ges-1_{pr::gfp}^{mt}* worms³⁹. The worms were synchronized by bleaching and allowed to hatch on plates containing *P. aeruginosa* and raised for 48 h at 25°C. Visualization of *hsp-6_{pr::gfp}*, *hsp-60_{pr::gfp}* and *atfs-1_{pr::atfs-1::gfp}* was performed essentially as described^{36,37}. *P. aeruginosa* was grown at 16°C for 24 h and seeded onto slow-killing plates. Plates were incubated overnight at room temperature. Synchronized L1 animals were transferred to *P. aeruginosa* plates and incubated at 20°C for 24 h before imaging.

To examine growth rates, eggs were allowed to hatch on plates containing *P. aeruginosa* and raised for 3 days at 25°C. 30 $\mu\text{g ml}^{-1}$ ethidium bromide or 0.2 mM paraquat was added to *E. coli* or *P. aeruginosa* slow-killing plates. For *clk-1(qm30)* growth rates, worms were raised for 4 days at 25°C.

Statistics. All experiments were performed three times yielding similar results and comprised of biological replicates. The sample size and statistical tests were chosen based on previous studies with similar methodologies and the data met the assumptions for each statistical test performed. No statistical methods were used in deciding sample sizes, nor were any blinded experiments performed. For all figures, the mean \pm standard deviation (s.d.) is represented unless otherwise noted.

C. elegans liquid-killing assay. *glp-4(bn2)* worms were raised at 25°C to sterilize them while being fed *atfs-1(RNAi)*. At the L4-early adult stage, the described worms were exposed to *P. aeruginosa* under conditions used for the liquid-killing assay⁴⁰.

RNA isolation and quantitative real-time PCR (qRT–PCR). Total RNA was obtained using the RNA STAT reagent (Tel-Test) and used for cDNA synthesis via the iScript cDNA Synthesis Kit (Bio-Rad Laboratories). qRT–PCR was performed using Thermo-Scientific SyBr Green Maxima Mix. For Fig. 1c–f, l, worms were hatched onto RNAi-expressing plates and harvested after 48 h. For Figs 2e–h, synchronized L4 worms were fed on *E. coli* or *P. aeruginosa* for 8 h using the slow-killing method before sample collection. For Extended Data Fig. 2a, b, synchronized *glp-4(bn2)* L4 worms were raised in liquid culture using *E. coli* or *P. aeruginosa* for 16 h. All values were normalized to wild-type worms grown on control bacteria for RNAi experiments (Fig. 1c–f, l) or wild-type worms grown on *E. coli* for *P. aeruginosa* experiments (Fig. 2e–h). *act-3* and *snb-1* mRNA were used as controls for slow-killing and liquid-killing experiments, respectively. *HPRT* mRNA was used as a control for dominant-negative AFG3L2 and ΔOTC experiments.

Primer sequences used for qRT–PCR were *act-3*: forward ATCCGTAAGGA CTTGTACGCCAAC and reverse CGATGATCTTGATCTTCATGGTTC; *abf-2*:

forward CGTGGCTGCCGACATCGACTT and reverse ATGCACAACCCCTGAG CCGC; *lys-2*: forward ATCGACTCGAACCAAGCTGCG and reverse TCGACA GCATTTCCCATTTGAAGCGT; *dec-4*: forward GAGCGACACTGGTGACTGTG and reverse CCATCCAGAATAGGTTGGCG; *clec-65*: forward CCCGGTGGTGGA CTGTGAATA and reverse AGCTCATATTGTGCTGGCA; *zip-2*: forward TCG ACGAGCAAACGACCTAC and reverse CTTGTGGCGTGCTCATGTT; *hsp-60*: forward AGGGATTTCGAGAGCATTCTGTCAG and reverse TGTGGCGACTT GAGCGATCTCTTC; *hsp-6*: forward GAAGATACGAAGACCCAGAGGTTTC and reverse CAACCTGAGATGGGGAATACACT; *snb-1*: forward CCGGATAAGA CCATCTTGACG and reverse GACGACTTCATCAACCTGAGC; *hBD-2*: forward GCCTCTTCCAGGTGTTTTC and reverse GAGACCACAGGTGCCAATTT; *hBD-4*: forward ATGTGGTTATGGGACTGCC and reverse AGCATGCATAG GTGTTGGGA; *HD-5*: forward TCCTTGCTGCCATTCTCCTG and reverse AC TGCTTCTGGGTTGTAGCC; *LL-37*: forward GCTGGGTGATTTCTCCGGA and reverse CCTGGGTACAAGATTCCGCA; *HPRT*: forward CTTTGCTGAC CTGCTGGATT and reverse TCCCTGTTGACTGGTCATT.

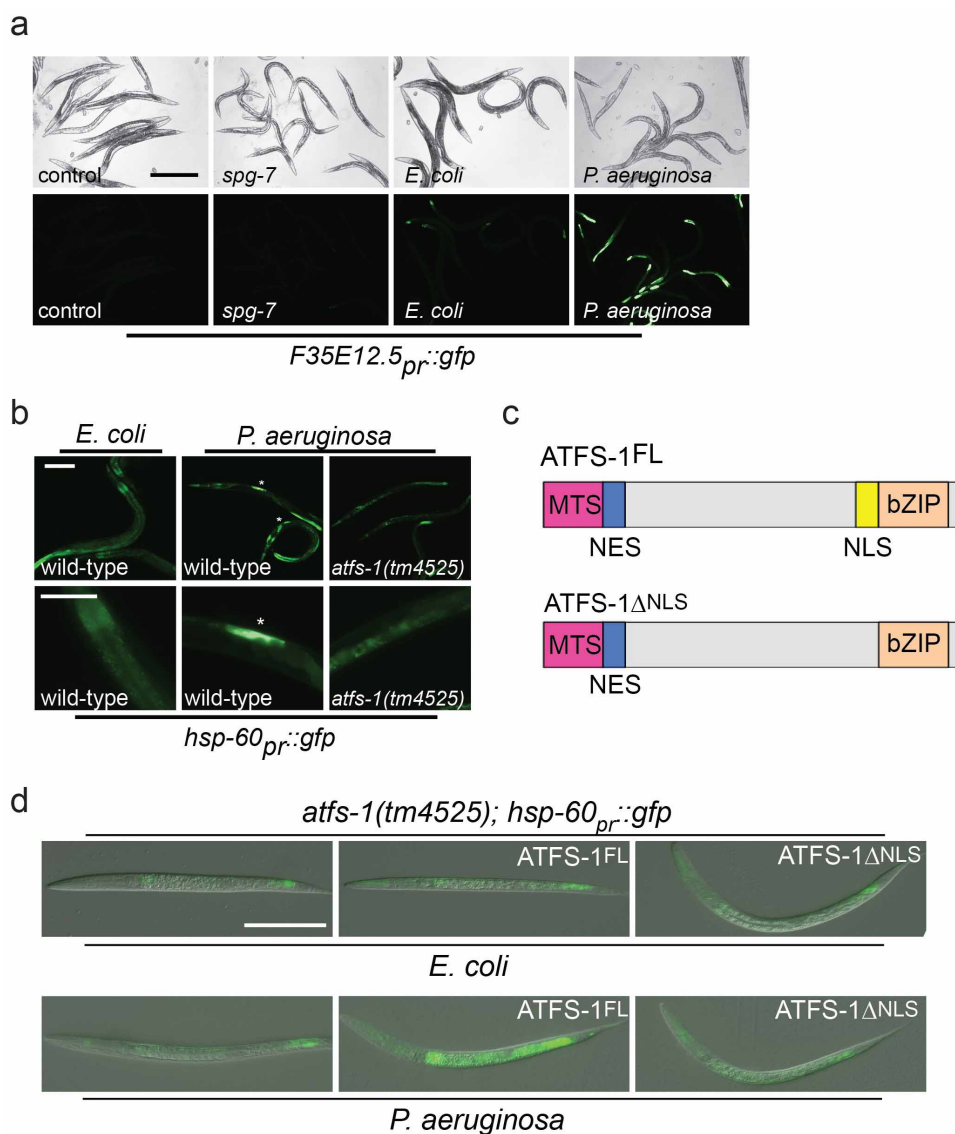
P. aeruginosa intestinal accumulation assay. To examine bacterial accumulation in the worm intestine, wild-type or *atfs-1(tm4919)* worms were synchronized and raised on control or *spg-7(RNAi)* plates for 48 h. Overnight cultures of *P. aeruginosa* expressing GFP (*P. aeruginosa*–GFP) were seeded onto slow-killing NGM plates, allowed to dry overnight at room temperature and then incubated at 37°C for 24 h. To exclude pathogen avoidance as a means of decreased intestinal colonization, where indicated *P. aeruginosa*–GFP was also spread across the entire surface of the slow-killing plate (Extended Data Fig. 5d, e). Worms at the L4 stage were transferred to *P. aeruginosa*–GFP plates and allowed to feed for 24–48 h before examination. The extent of bacterial accumulation was scored as either ‘none/mild’, ‘moderate’ or ‘strong’ as indicated (Extended Data Fig. 5c).

Plasmid construction. The *hsp-16_{pr::atfs-1^{FL}}* and *hsp-16_{pr::atfs-1^{ANLS}}* plasmids were described previously³⁷. To construct the *lys-2_{pr::gfp}* plasmid, a 803 base pair fragment of the *lys-2* promoter sequence upstream of the start codon was amplified using PCR and cloned into the HindIII and PstI sites of pPD95.75. *lys-2_{pr::gfp}* was microinjected into wild-type worms at a concentration of 20 ng μl^{-1} along with *myo-3_{pr::mCherry}* at a concentration of 60 ng μl^{-1} .

P. aeruginosa avoidance assay. Synchronized L1 wild-type and *atfs-1(tm4919)* worms were allowed to develop on control or *spg-7(RNAi)* plates to the L4 stage and then transferred to *E. coli* or *P. aeruginosa* slow-killing plates for 17 h when the worms were scored. The extent of avoidance was expressed as the per cent of animals off of the bacterial lawn over the total number of animals on the plate (Extended Data Fig. 5a, b).

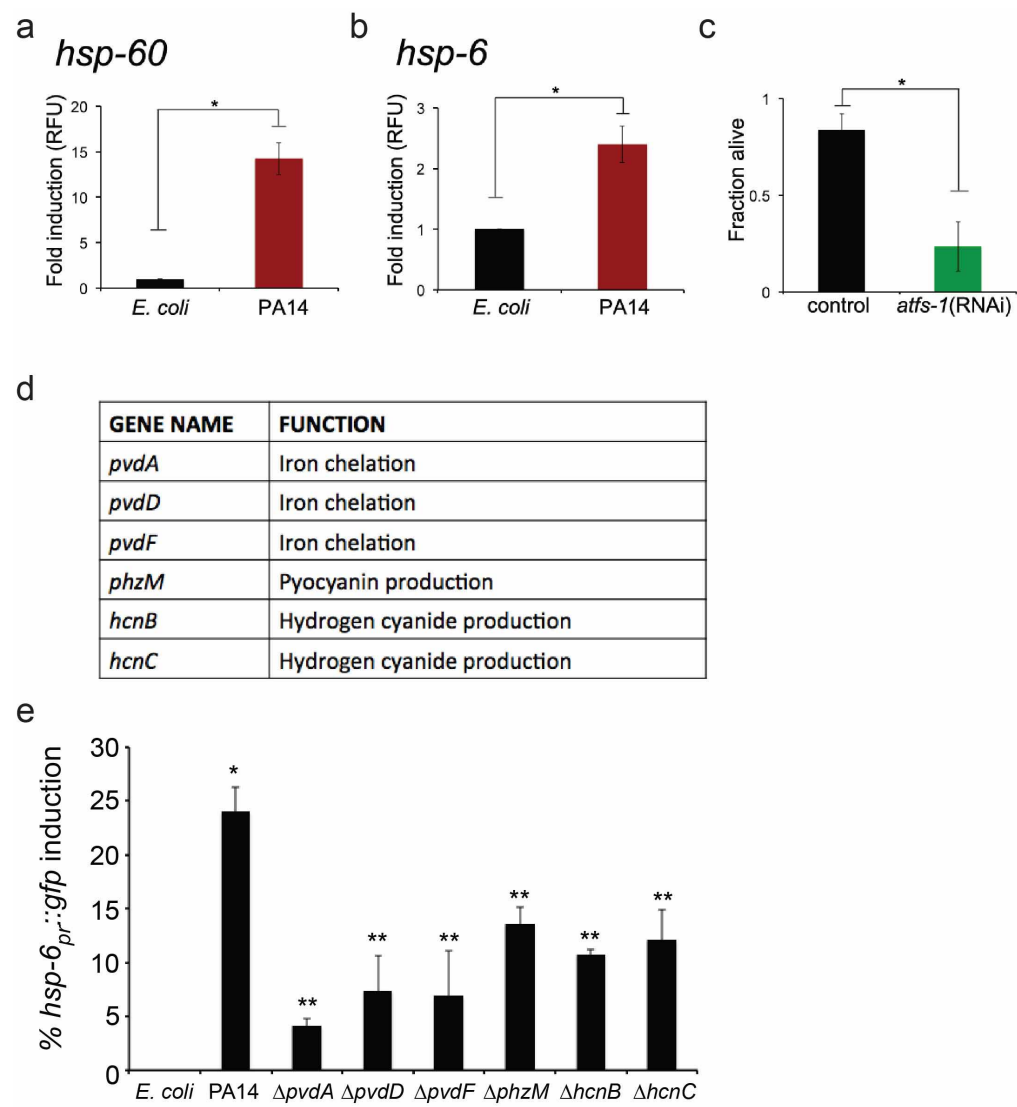
Microscopy. *C. elegans* were imaged using a Zeiss AxioCam MRm mounted on a Zeiss Imager.Z2 microscope. Exposure times were the same in each experiment.

- Yoneda, T. *et al.* Compartment-specific perturbation of protein handling activates genes encoding mitochondrial chaperones. *J. Cell Sci.* **117**, 4055–4066 (2004).
- Rual, J. F. *et al.* Toward improving *Caenorhabditis elegans* genome mapping with an ORFeome-based RNAi library. *Genome Res.* **14**, 2162–2168 (2004).
- Ehres, S. *et al.* Regulation of OPA1 processing and mitochondrial fusion by *m*-AAA protease isoenzymes and OMA1. *J. Cell Biol.* **187**, 1023–1036 (2009).
- Zhao, Q. *et al.* A mitochondrial specific stress response in mammalian cells. *EMBO J.* **21**, 4411–4419 (2002).
- Tan, M. W., Mahajan-Miklos, S. & Ausubel, F. M. Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. *Proc. Natl Acad. Sci. USA* **96**, 715–720 (1999).
- Liu, Y., Samuel, B. S., Breen, P. C. & Ruvkun, G. *Caenorhabditis elegans* pathways that surveil and defend mitochondria. *Nature* **508**, 406–410 (2014).
- Nargund, A. M., Pellegrino, M. W., Fiorese, C. J., Baker, B. M. & Haynes, C. M. Mitochondrial import efficiency of ATFS-1 regulates mitochondrial UPR activation. *Science* **337**, 587–590 (2012).
- Kennedy, S., Wang, D. & Ruvkun, G. A conserved siRNA-degrading RNase negatively regulates RNA interference in *C. elegans*. *Nature* **427**, 645–649 (2004).
- Benedetti, C., Haynes, C. M., Yang, Y., Harding, H. P. & Ron, D. Ubiquitin-like protein 5 positively regulates chaperone gene expression in the mitochondrial unfolded protein response. *Genetics* **174**, 229–239 (2006).
- Kirienko, N. V. *et al.* *Pseudomonas aeruginosa* disrupts *Caenorhabditis elegans* iron homeostasis, causing a hypoxic response and death. *Cell Host Microbe* **13**, 406–416 (2013).



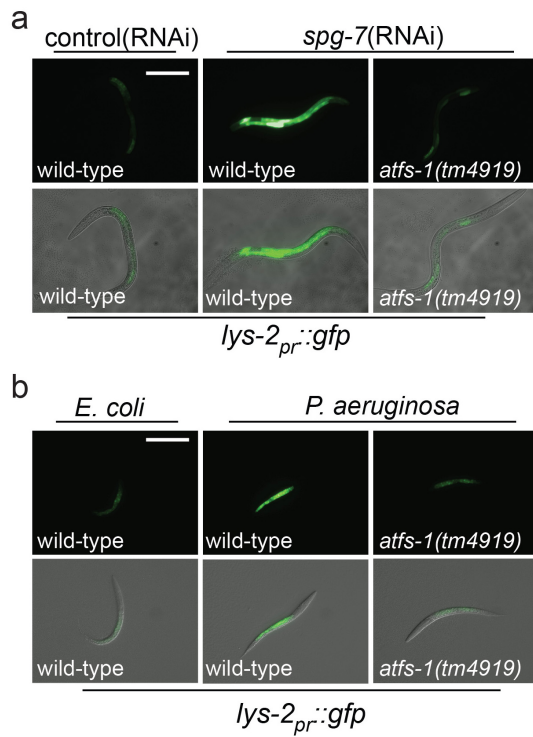
Extended Data Figure 1 | Nuclear accumulation of ATFS-1 is required for UPR^{mt} activation during *P. aeruginosa* exposure. **a**, Representative photomicrographs of *F35E12.5_{pr}::gfp* transgenic worms raised on control or *spg-7*(RNAi). No detectable increase in expression was observed following *spg-7*(RNAi) treatment. In contrast, strong expression of *F35E12.5_{pr}::gfp* was observed following exposure to *P. aeruginosa* compared to *E. coli* controls. Scale bar, 0.5 mm. **b**, Wild-type or *atfs-1(tm4525);hsp-60_{pr}::gfp* worms on *E. coli* or

P. aeruginosa. Lower panels are magnified views of the intestine showing enhanced expression of *hsp-60_{pr}::gfp* (asterisks). Scale bars, 0.05 mm. **c**, Diagrams of wild-type ATFS-1 (ATFS-1^{FL}) and ATFS-1 with a mutated nuclear localization signal (ATFS-1^{ΔNLS}). **d**, Photomicrographs of *atfs-1(tm4525);hsp-60_{pr}::gfp* worms expressing ATFS-1^{FL} or ATFS-1^{ΔNLS} via the *hsp-16* promoter exposed to *E. coli* or *P. aeruginosa*. Scale bar, 0.1 mm.



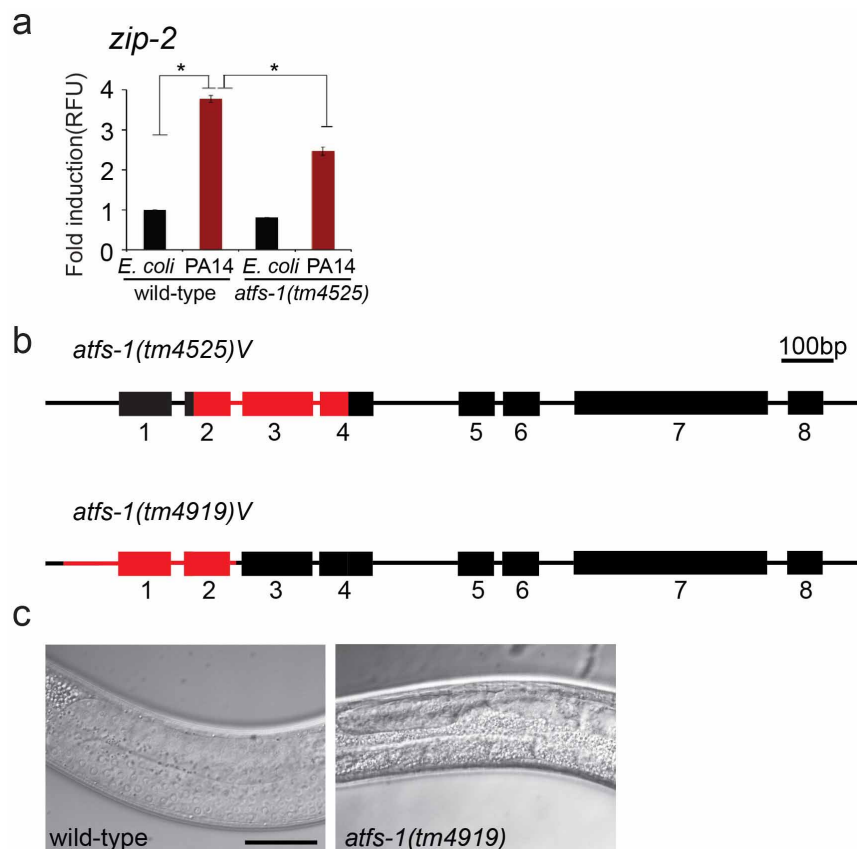
Extended Data Figure 2 | Multiple *P. aeruginosa* virulence genes contribute to UPR^{mt} activation. **a, b**, Expression of *hsp-60* and *hsp-6* mRNA for *glp-4(bn2)* worms exposed to *E. coli* or *P. aeruginosa* liquid-killing using qRT-PCR ($n = 3$, \pm s.d.). Fold inductions are normalized to wild-type *E. coli* test group, $*P < 0.05$ (Student's *t*-test). **c**, Quantification of survival for *glp-4(bn2)* worms raised on control or *atfs-1*(RNAi) and exposed to *P. aeruginosa* liquid-killing, $*P < 0.0001$ (Student's *t*-test). **d**, List of *P. aeruginosa* toxin mutants.

e, Quantification of the proportion of worms showing increased *hsp-6_{pr}::gfp* expression in the intestine under slow-killing conditions. Exposure to *P. aeruginosa* caused *hsp-6_{pr}::gfp* induction ($n = 3$, \pm s.e.m.), $*P < 0.05$ (Student's *t*-test). However, exposure to *P. aeruginosa* with mutations in the *pvdA*, *pvdD*, *pvdF*, *phzM*, *hcnB*, or *hcnC* toxin genes resulted in relatively less UPR^{mt} activation ($n = 3$, \pm s.e.m.), $**P < 0.05$ (Student *t*-test).



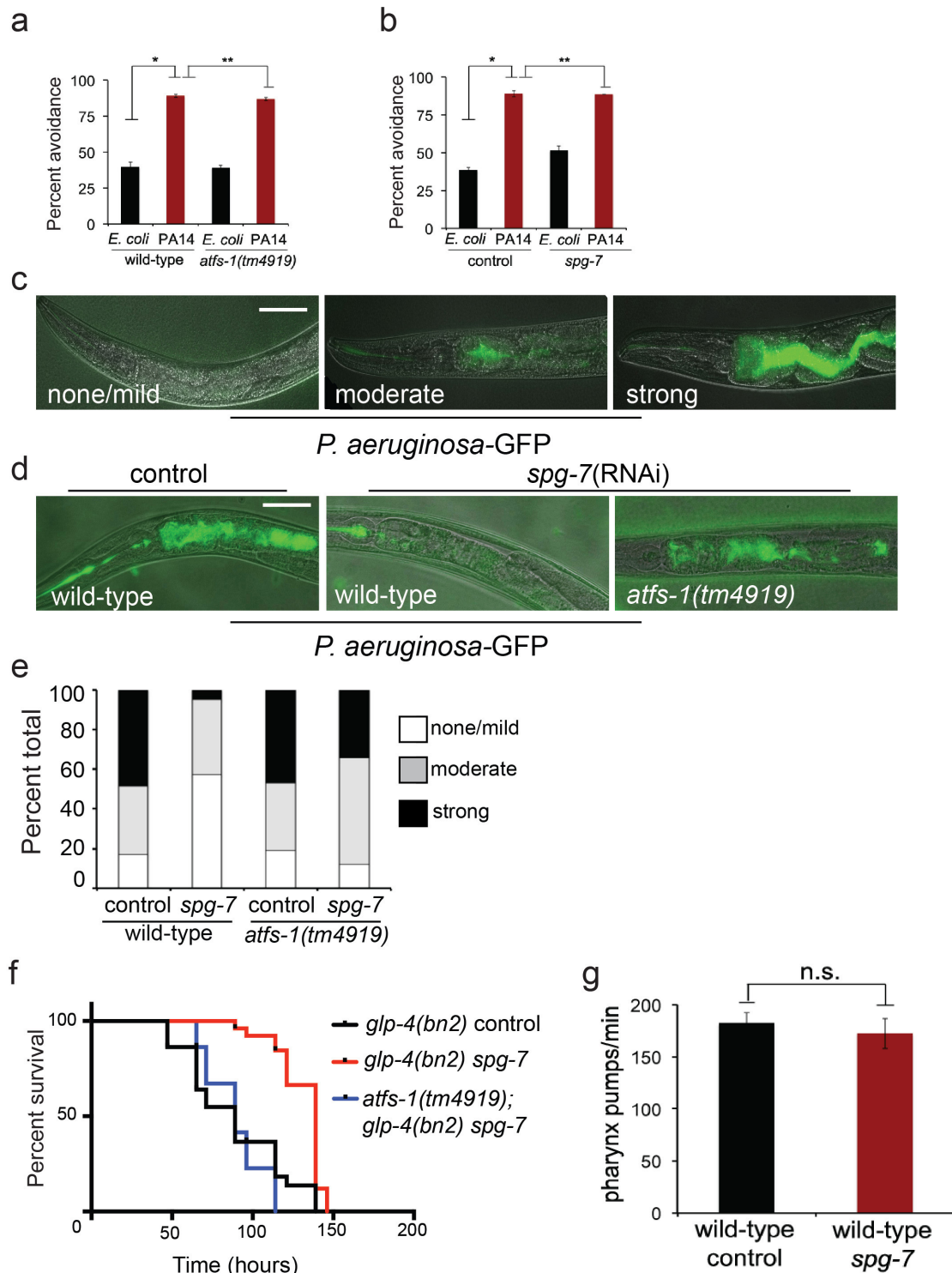
Extended Data Figure 3 | Intestinal accumulation of *lys-2* during mitochondrial stress and *P. aeruginosa* exposure requires ATFS-1.

a, Representative photomicrographs of wild-type and *atfs-1(tm4919)* worms carrying the *lys-2_{pr}::gfp* transgene raised on control or *spg-7*(RNAi). Scale bar, 0.1 mm. **b**, Representative photomicrographs of wild-type and *atfs-1(tm4919)* worms carrying the *lys-2_{pr}::gfp* transgene exposed to *E. coli* or *P. aeruginosa*. Scale bar, 0.1 mm.



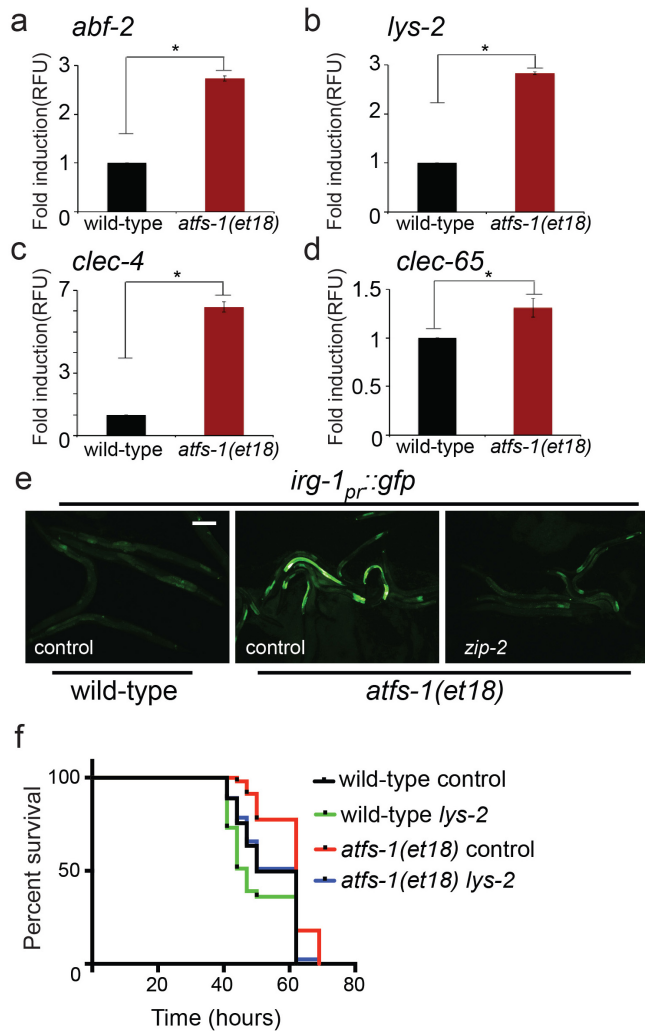
Extended Data Figure 4 | ATFS-1 partially regulates *zip-2* expression during *P. aeruginosa* exposure. **a**, Expression levels of *zip-2* mRNA in wild-type or *atfs-1(tm4525)* worms raised on *E. coli* or *P. aeruginosa* using qRT-PCR ($n = 3, \pm$ s.d.), * $P < 0.05$ (Student's t test). **b**, Schematic diagram of the *atfs-1* genomic open reading frame showing positions of exons 1–8 (boxes) and

locations of the *tm4525* (ref. 5) and *tm4919* deletions in red. The *tm4919* allele is a 334 base pair deletion beginning 107 base pairs upstream of the *atfs-1* start codon and ends within the second intron of the *atfs-1* genomic open reading frame. **c**, Representative photomicrographs of a germline in wild-type and *atfs-1(tm4919)* worms. Scale bar, 0.02 mm.

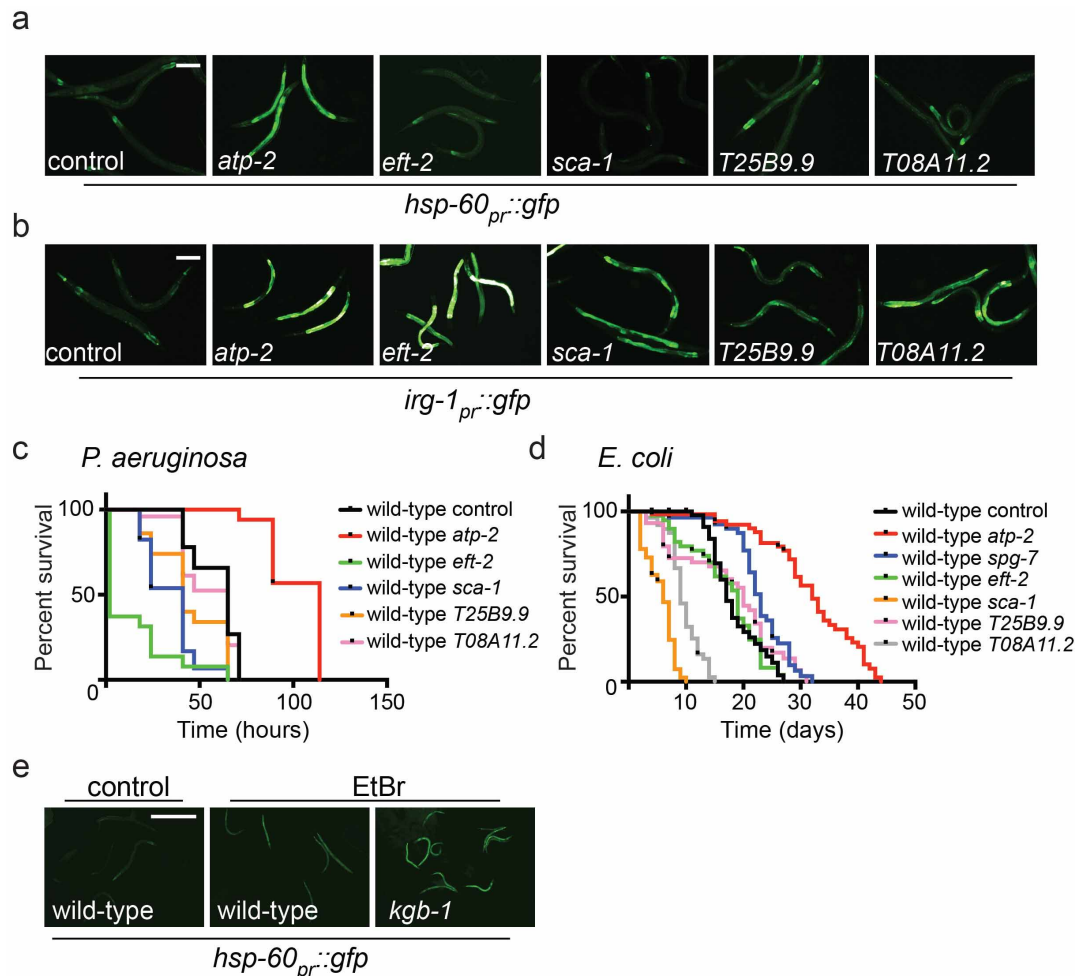


Extended Data Figure 5 | ATFS-1 is not required for pathogen avoidance during *P. aeruginosa* exposure. **a**, Quantification of avoidance behaviour for wild-type and *atfs-1(tm4919)* worms raised on *E. coli* or *P. aeruginosa*, expressed as a percentage of the number of animals off the bacterial lawn relative to the total number of worms ($n = 4$, \pm s.d.). $*P < 0.0001$, $**P = 0.1914$ (Student's *t*-test). **b**, Quantification of avoidance behaviour for wild-type worms raised on control or *spg-7*(RNAi) and exposed to *E. coli* or *P. aeruginosa*, expressed as a percentage of the number of animals off the bacterial lawn relative to the total number of worms ($n = 3$, \pm s.d.). $*P < 0.0001$, $**P = 0.8706$ (Student's *t*-test). **c**, Representative photomicrographs illustrating the scored level of infection for *P. aeruginosa* colonization assay using *P. aeruginosa*-GFP. Three categories of *P. aeruginosa*-GFP infection were used: none/mild, moderate and strong. Scale bar, 0.1 mm.

d, Representative photomicrographs of wild-type and *atfs-1(tm4919)* worms raised on *spg-7*(RNAi) and exposed to a lawn of *P. aeruginosa*-GFP that completely covered the surface of the slow-killing plate for 24 h. Images are overlays of DIC and GFP. Scale bar, 0.1 mm. **e**, Quantification of *P. aeruginosa* intestinal colonization as shown in Extended Data Fig. 5d. White, grey and black bars denote no/mild infection, moderate infection and strong infection, respectively. Forty worms were analysed per treatment. **f**, Survival analysis of *glp-4(bn2)* and *atfs-1(tm4919); glp-4(bn2)* worms raised on control or *spg-7*(RNAi) and exposed to *P. aeruginosa*. Statistics for each survival analysis are presented in Extended Data Table 3. **g**, Quantification of pharyngeal pumping rate per minute for wild-type worms raised on control or *spg-7*(RNAi) ($n = 10$, \pm s.d.). n.s., no significant difference ($P = 0.10$; Student's *t*-test).



Extended Data Figure 6 | *atfs-1(et18)* gain of function mutant worms induce innate immune gene expression in the absence of mitochondrial stress. **a–d**, Expression levels of *abf-2*, *lys-2*, *clec-4* and *clec-65* mRNA in wild-type or *atfs-1(et18)* worms using qRT-PCR ($n = 3$, \pm s.d.), $*P < 0.05$ (Student's t test). **e**, Representative photomicrographs of wild-type and *atfs-1(et18)* worms carrying the *irg-1_{pr}::gfp* transgene raised on control or *zip-2*(RNAi). Scale bar, 0.10 mm. **f**, Survival analysis of wild-type and *atfs-1(et18)* worms raised on control or *lys-2*(RNAi) and exposed to *P. aeruginosa*. Statistics for each survival analysis are presented in Extended Data Table 3.



Extended Data Figure 7 | Mitochondrial protective and innate immune gene induction contributes to ATFS-1-mediated resistance to *P. aeruginosa* infection. **a**, Representative photomicrographs of wild-type *hsp-60_{pr}::gfp* worms raised on control, *atp-2*(RNAi), *spg-7*(RNAi), *eft-2*(RNAi), *sca-1*(RNAi), T25B9.9(RNAi) or T08A11.2(RNAi). Scale bar is 0.1 mm. **b**, Representative photomicrographs of wild-type *irg-1_{pr}::gfp* worms raised on control, *atp-2*(RNAi), *eft-2*(RNAi), *sca-1*(RNAi), T25B9.9(RNAi) or T08A11.2(RNAi). Scale bar is 0.1 mm. **c**, Survival analysis of wild-type worms raised on control, *atp-2*(RNAi), *eft-2*(RNAi), *sca-1*(RNAi), T25B9.9(RNAi) or

T08A11.2(RNAi) and exposed to *P. aeruginosa*. Statistics for each survival analysis are presented in Extended Data Table 3. **d**, Survival analysis of wild-type worms raised on control, *atp-2*(RNAi), *spg-7*(RNAi), *eft-2*(RNAi), T25B9.9(RNAi) or T08A11.2(RNAi) and exposed to *E. coli*. Statistics for each survival analysis are presented in Extended Data Table 3. **e**, Representative photomicrographs of wild-type or *kbg-1(km21);hsp-60_{pr}::gfp* worms raised on *E. coli* plates with or without 30 $\mu\text{g ml}^{-1}$ ethidium bromide, suggesting that the KGB-1 Jun kinase pathway negatively regulates the UPR^{mt} during mitochondrial stress²⁹. Scale bar, 0.5 mm.

Extended Data Table 1 | ATFS-1-dependent innate immune genes upregulated when raised on *spg-7*(RNAi)⁵

Sequence Name	Gene symbol	KOG title, protein domain or function	Fold Induction Wild-type <i>spg-7</i> (RNAi)/control	Fold Induction <i>atfs-1(tm4525)</i> <i>spg-7</i> (RNAi)/control
<i>Antimicrobial peptides</i>				
g6714550	<i>abf-2</i>	antimicrobial peptide	10.456	2.33072
R09B5.9	<i>cnc-4</i>	Caenorhabditis bacteriocin	4.58689	1.59881
<i>Lysozyme</i>				
Y22F5A.5	<i>lys-2</i>	N-acetylmuraminidase/lysozyme	4.81394	2.87725
<i>C-type lectins</i>				
F35C5.9	<i>clec-66</i>	Lectin C-type domain/CUB domain	3.60596	2.09418
F35C5.5	<i>clec-62</i>	Lectin C-type domain/CUB domain	3.81898	2.47423
F35C5.8	<i>clec-65</i>	Lectin C-type domain	4.366	2.25587
E03H4.10	<i>clec-17</i>	C-type lectin	21.9148	7.58123
C03H5.1	<i>clec-10</i>	C-type lectin	3.47056	1.82945
F31D4.4	<i>clec-264</i>	C-type lectin	1.87899	1.24332
T09F5.9	<i>clec-47</i>	C-type lectin	5.57165	-1.12101
Y38E10A.5	<i>clec-4</i>	C-type lectin	8.25168	3.66069
M02F4.7	<i>clec-265</i>	C-type lectin	8.40775	4.75704
F08H9.7	<i>clec-56</i>	C-type lectin	2.30962	1.32766
<i>Galectin</i>				
F38A5.3	<i>lec-11</i>	Galectin, galactose-binding lectin	2.29608	1.41767
<i>Signaling</i>				
F08B1.1	<i>vhp-1</i>	Dual specificity phosphatase	2.39868	1.59839

Extended Data Table 2 | ATFS-1-dependent UPR^{mt} genes in common with genes induced following *P. aeruginosa* exposure¹⁰

Sequence name	Gene symbol	KOG title, protein domain or function	Fold Induction	Fold Induction
			Wild-type <i>spg-7</i> (RNAi)/control (Nargund et al., 2012)	Wild-type PA14/OP50 (Troemel et al. 2006 or this study*)
R08F11.3	<i>cyp-33C8</i>	Cytochrome P450 CYP2 subfamily	52.7502	4.1
T10B9.2	<i>cyp-13A5</i>	Cytochrome P450 family	35.6897	2.8
E03H4.10	<i>clec-17</i>	C-type lectin	21.9148	7.9
C54D10.1	<i>cdr-2</i>	glutathione S-transferase-like protein	15.2001	2.6
K01D12.11	<i>cdr-4</i>	cadmium responsive	11.9017	4
g6714550	<i>abf-2</i>	antimicrobial peptide	10.456	1.89*
F15B9.6			9.37816	2.7
K10D11.2			8.84434	3.1
M02F4.7	<i>clec-265</i>	C-type Lectin	8.40775	3.9
Y38E10A.5	<i>clec-4</i>	C-type Lectin	8.25168	11.1
C49G7.5			7.7102	21
C18A11.1			7.43593	3.1
F22H10.2			7.33184	4.4
C49G7.10			7.14237	9.7
C10C5.2			6.85214	5.9
Y58A7A.5			5.84559	12
R11G11.12	<i>nhr-210</i>	Nuclear Hormone Receptor family	5.76655	2.3
T12G3.1		contains ZZ-type Zn-finger	5.74038	3
ZK970.7			5.46349	3.9
R09B5.9	<i>cnc-4</i>	Caenorhabditis bacteriocin	4.58689	4.2
C49G7.7			4.41335	5.6
F35C5.8	<i>clec-65</i>	Lectin C-type domain	4.366	2.9
Y58A7A.3			4.34718	5.7
C34H4.2			4.34064	2.3
T16G1.5		Predicted small molecule kinase	4.25188	8.9
		Hydroxyacyl-CoA dehydrogenase/enoyl-CoA hydratase		5
F01G10.3	<i>ech-9</i>	CoA hydratase	4.17949	
T12G3.1		Uncharacterized conserved protein	4.09838	2.4
C50F4.1			4.07802	2.1
B0218.2	<i>faah-2</i>	amidase	3.8877	2.2
F19B2.5		Helicase-like transcription factor	3.81204	2.3
C50F4.1			3.78737	2.1
F35C5.9	<i>clec-66</i>	Lectin C-type domain/CUB domain	3.60596	6.2
M01G12.9			3.57904	3.2
K10G4.3			3.56853	2.2
C03H5.1	<i>clec-10</i>	C-type lectin	3.47056	2.3
C34C6.7			3.17263	2.6
Y22D7AR.9	<i>fbxa-74</i>	F-box A protein	3.10471	8.7
Y119D3B.20	<i>fbxa-92</i>	F-box A protein	3.06775	4.1
F55C12.7	<i>tag-234</i>		3.02488	3.5
Y17G7B.8			3.02481	2.2
C29F9.3			3.02333	2.6
C34D1.5	<i>zip-5</i>	bZip transcription factor	2.98077	2
Y58A7A.4			2.95958	2.2
T01D3.6		von Willebrand factor	2.91981	2.6
ZK418.7			2.55033	2.4
F49F1.6		Secreted surface protein	2.48376	17.6
K08D8.6			2.41167	3.6
C09F12.1	<i>clc-1</i>	claudin homolog	2.36493	2.3
T27F2.4		bZip transcription factor	2.35058	2.8
F38A5.3	<i>lec-11</i>	Galectin, galactose-binding lectin	2.29608	2.4
Y47H10A.5			2.24131	3.9
Y43C5A.3			2.06783	2.5
F23H12.3			1.77139	2
E02C12.8		Predicted small molecule kinase	1.7588	3.7
Y95B8A.6			1.67149	2.5
F11D11.3			1.55814	3.4
C50F4.9			1.54636	3.4
Y22F5A.5	<i>lys-2</i>	lysozyme	4.81394	10.6*

*this study

Extended Data Table 3 | Statistics for survival analysis

Strain comparison	p values	Number of worms	Figure
<i>eri(mg366)</i> control vs <i>eri-1(mg366) affs-1(RNAi)</i>	0.0003	<i>eri(mg366)</i> control: 34/50, <i>eri-1(mg366) affs-1(RNAi)</i> : 39/50	3a
<i>eri(mg366)</i> control vs <i>eri-1(mg366) affs-1(RNAi)</i>	0.6986	<i>eri(mg366)</i> control: 50/50, <i>eri-1(mg366) affs-1(RNAi)</i> : 47/50	3b
wild-type control vs wild-type <i>spg-7(RNAi)</i>	<0.0001	wild-type control: 37/50, wild-type <i>spg-7(RNAi)</i> : 45/50	3e
wild-type <i>spg-7(RNAi)</i> vs <i>affs-1(tm4919) spg-7(RNAi)</i>	<0.0001	wild-type <i>spg-7(RNAi)</i> : 37/50, <i>affs-1(tm4919) spg-7(RNAi)</i> : 50/50	3e
wild-type control vs <i>affs-1(tm4919) spg-7(RNAi)</i>	0.1322	wild-type control: 37/50, <i>affs-1(tm4919) spg-7(RNAi)</i> : 50/50	3e
wild-type control vs <i>affs-1(et18)</i> control	<0.0001	wild-type control: 41/50, <i>affs-1(et18)</i> control: 30/50	3h
wild-type control vs wild-type <i>affs-1(RNAi)</i>	0.039	wild-type control: 41/50, wild-type <i>affs-1(RNAi)</i> : 45/50	3h
<i>affs-1(et18)</i> control vs <i>affs-1(et18) affs-1(RNAi)</i>	<0.0001	<i>affs-1(et18)</i> control: 30/50, <i>affs-1(et18) affs-1(RNAi)</i> : 37/50	3h
wild-type control vs <i>pmk-1(km25)</i> control	<0.0001	wild-type control: 33/50, <i>pmk-1(km25)</i> control: 42/50	4a
wild-type control vs wild-type <i>spg-7(RNAi)</i>	<0.0001	wild-type control: 33/50, wild-type <i>spg-7(RNAi)</i> : 45/50	4a
<i>pmk-1(km25)</i> control vs <i>pmk-1(km25) spg-7(RNAi)</i>	<0.0001	<i>pmk-1(km25)</i> control: 42/50, <i>pmk-1(km25) spg-7(RNAi)</i> : 23/50	4a
wild-type control vs <i>sek-1(km4)</i> control	<0.0001	wild-type control: 35/50, <i>sek-1(km4)</i> control: 35/50	4b
wild-type control vs wild-type <i>spg-7(RNAi)</i>	<0.0001	wild-type control: 35/50, wild-type <i>spg-7(RNAi)</i> : 37/50	4b
<i>sek-1(km4)</i> control vs <i>sek-1(km4) spg-7(RNAi)</i>	<0.0001	<i>sek-1(km4)</i> control: 35/50, <i>sek-1(km4) spg-7(RNAi)</i> : 39/50	4b
wild-type control vs <i>kgb-1(km21)</i> control	<0.0001	wild-type control: 29/50, <i>kgb-1(km21)</i> control: 38/50	4c
wild-type control vs wild-type <i>spg-7(RNAi)</i>	<0.0001	wild-type control: 29/50, wild-type <i>spg-7(RNAi)</i> : 27/50	4c
<i>kgb-1(km21)</i> control vs <i>kgb-1(km21) spg-7(RNAi)</i>	<0.0001	<i>kgb-1(km21)</i> control: 38/50, <i>kgb-1(km21) spg-7(RNAi)</i> : 29/50	4c
wild-type control vs <i>mlk-1(ok2471)</i> control	0.0001	wild-type control: 38/50, <i>mlk-1(ok2471)</i> control: 28/50	4d
wild-type control vs wild-type <i>spg-7(RNAi)</i>	<0.0001	wild-type control: 38/50, wild-type <i>spg-7(RNAi)</i> : 33/50	4d
<i>mlk-1(ok2471)</i> control vs <i>mlk-1(ok2471) spg-7(RNAi)</i>	<0.0001	<i>mlk-1(ok2471)</i> control 28/50 vs <i>mlk-1(ok2471) spg-7(RNAi)</i> : 49/50	4d
wild-type control vs wild-type <i>spg-7(RNAi)</i>	<0.0001	wild-type control: 50/50, wild-type <i>spg-7(RNAi)</i> : 28/50	4e
wild-type <i>spg-7(RNAi)</i> vs <i>affs-1(tm4919) spg-7(RNAi)</i>	<0.0001	wild-type <i>spg-7(RNAi)</i> : 28/50, <i>affs-1(tm4919) spg-7(RNAi)</i> : 28/50	4e
wild-type <i>spg-7(RNAi)</i> vs <i>zip-2(4248) spg-7(RNAi)</i>	0.0004	wild-type <i>spg-7(RNAi)</i> : 28/50, <i>zip-2(4248) spg-7(RNAi)</i> : 39/50	4e
<i>glp-4(bn2)</i> control vs <i>glp-4(bn2) spg-7(RNAi)</i>	<0.0001	<i>glp-4(bn2)</i> control: 32/50, <i>glp-4(bn2) spg-7(RNAi)</i> : 19/50	ED 5f
<i>glp-4(bn2) spg-7(RNAi)</i> vs <i>affs-1(tm4919) spg-7(RNAi)</i>	<0.0001	<i>glp-4(bn2) spg-7(RNAi)</i> : 19/50, <i>affs-1(tm4919) spg-7(RNAi)</i> : 32/50	ED 5f
wild-type control vs <i>affs-1(et18)</i> control	0.0001	wild-type control: 40/50, <i>affs-1(et18)</i> control: 35/50	ED 6f
wild-type control vs wild-type <i>lys-2(RNAi)</i>	0.0419	wild-type control: 40/50, wild-type <i>lys-2(RNAi)</i> : 38/50	ED 6f
<i>affs-1(et18)</i> control vs <i>affs-1(et18) lys-2(RNAi)</i>	0.0004	<i>affs-1(et18)</i> control: 35/50, <i>affs-1(et18) lys-2(RNAi)</i> : 44/50	ED 6f
wild-type control vs <i>affs-1(et18) lys-2(RNAi)</i>	0.7317	wild-type control: 40/50, <i>affs-1(et18) lys-2(RNAi)</i> : 44/50	ED 6f
wild-type control vs wild-type <i>atp-2</i>	<0.0001	wild-type control: 36/50, wild-type <i>atp-2</i> : 31/50	ED 7c
wild-type control vs wild-type <i>eff-2</i>	<0.0001	wild-type control: 36/50, wild-type <i>eff-2</i> : 50/50	ED 7c
wild-type control vs wild-type <i>sca-1</i>	<0.0001	wild-type control: 36/50, wild-type <i>sca-1</i> : 34/50	ED 7c
wild-type control vs wild-type T25B9.9	<0.0001	wild-type control vs wild-type T25B9.9: 50/50	ED 7c
wild-type control vs wild-type T08A11.2	0.1536	wild-type control vs wild-type T08A11.2: 37/50	ED 7c
wild-type control vs wild-type <i>atp-2</i>	<0.0001	wild-type control: 37/50, wild-type <i>atp-2</i> : 31/50	ED 7d
wild-type control vs wild-type <i>spg-7(RNAi)</i>	<0.0001	wild-type control: 37/50, wild-type <i>spg-7(RNAi)</i> : 38/50	ED 7d
wild-type control vs wild-type <i>eff-2</i>	0.6435	wild-type control: 37/50, wild-type <i>eff-2</i> : 28/50	ED 7d
wild-type control vs wild-type <i>sca-1</i>	<0.0001	wild-type control: 37/50, wild-type <i>sca-1</i> : 48/50	ED 7d
wild-type control vs wild-type T25B9.9	0.2626	wild-type control: 37/50, wild-type T25B9.9: 38/50	ED 7d
wild-type control vs wild-type T08A11.2	<0.0001	wild-type control: 37/50, wild-type T08A11.2: 43/50	ED 7d

ED= Extended Data

Statistical analysis was performed using the log rank (Mantel-Cox) statistical test. Number of worms represents the number of dead worms scored relative to the number of worms alive at the start of the experiment. The difference in numbers indicates those worms that were excluded (see Methods).

Rapid development of broadly influenza neutralizing antibodies through redundant mutations

Leontios Pappas¹, Mathilde Foglierini¹, Luca Piccoli¹, Nicole L. Kallewaard², Filippo Turrini³, Chiara Silacci¹, Blanca Fernandez-Rodriguez¹, Gloria Agatic⁴, Isabella Giacchetto-Sasselli¹, Gabriele Pellicciotta⁵, Federica Sallusto¹, Qing Zhu², Elisa Vicenzi³, Davide Corti^{1,4*} & Antonio Lanzavecchia^{1,6*}

The neutralizing antibody response to influenza virus is dominated by antibodies that bind to the globular head of haemagglutinin, which undergoes a continuous antigenic drift, necessitating the re-formulation of influenza vaccines on an annual basis. Recently, several laboratories have described a new class of rare influenza-neutralizing antibodies that target a conserved site in the haemagglutinin stem^{1–6}. Most of these antibodies use the heavy-chain variable region *VH1-69* gene, and structural data demonstrate that they bind to the haemagglutinin stem through conserved heavy-chain complementarity determining region (HCDR) residues. However, the *VH1-69* antibodies are highly mutated and are produced by some but not all individuals^{6,7}, suggesting that several somatic mutations may be required for their development^{8,9}. To address this, here we characterize 197 anti-stem antibodies from a single donor, reconstruct the developmental pathways of several *VH1-69* clones and identify two key elements that are required for the initial development of most *VH1-69* antibodies: a polymorphic germline-encoded phenylalanine at position 54 and a conserved tyrosine at position 98 in HCDR3. Strikingly, in most cases a single proline to alanine mutation at position 52a in HCDR2 is sufficient to confer high affinity binding to the selecting H1 antigen, consistent with rapid affinity maturation. Surprisingly, additional favourable mutations continue to accumulate, increasing the breadth of reactivity and making both the initial mutations and phenylalanine at position 54 functionally redundant. These results define *VH1-69* allele polymorphism, rearrangement of the VDJ gene segments and single somatic mutations as the three requirements for generating broadly neutralizing *VH1-69* antibodies and reveal an unexpected redundancy in the affinity maturation process.

To understand the developmental pathways of broadly neutralizing influenza antibodies, we studied the antibody response against the haemagglutinin (HA) stem in a donor from whom we previously identified the pan-influenza A neutralizing antibody FI6 (ref. 4). Over a period of 5 years, we isolated 197 stem-specific antibodies from memory B cells or plasma cells following infection or vaccination. Forty per cent of the antibodies used the *VH1-69* gene while the remainder used other *VH* genes (Extended Data Fig. 1a). On the basis of unique VDJ and VJ rearrangements, we identified 17 clusters of clonally related antibodies, six of which use *VH1-69* (Supplementary Fig. 1). Among stem-specific antibodies, those using *VH1-69* carried the highest load of somatic mutations (mean \pm s.d. 29.8 ± 6.8 versus 22.6 ± 9.7). In contrast, antibodies to the globular head of the 2009 pandemic H1N1 virus isolated from the same individual carried significantly fewer somatic mutations (15.5 ± 7.3) (Extended Data Fig. 1b). These findings suggest that anti-stem antibodies, in particular those using *VH1-69*, arise from repeated stimulation in germinal centres.

To assess the contribution of somatic mutations to affinity maturation we compared the mutated and unmutated common ancestor (UCA)

versions of four *VH1-69* multi-member clones and 11 previously described antibodies using surface plasmon resonance (SPR) (Extended Data Fig. 1c and Supplementary Figs 2–4). With the exception of FI0, isolated from a random phage library¹, all UCA antibodies bound to soluble H1-HA. In most cases, mutated antibodies bound with increased affinity due to decreased dissociation rates, while association rates remained relatively constant (Supplementary Fig. 5). A comparison of antibodies formed by mutated *VH1-69* heavy chains paired with mutated, unmutated or irrelevant light chains confirmed that HA binding is exclusively mediated by the heavy chain^{1–3} (Extended Data Fig. 2). These findings are consistent with low-affinity binding of naive B cell receptors to the selecting H1-HA through heavy-chain residues only.

To identify the early steps in the development of high affinity anti-stem *VH1-69* antibodies, we used the classical approach of reconstructing genealogical trees of multimember clones^{10,11} using dnaml (PHYMLIP package) to infer the maximum likelihood trees¹² (Fig. 1a, b). When applied to clone 9, this analysis revealed an early branchpoint (BP9-1) containing two amino-acid substitutions (S30R/P52aA) shared by all five antibodies of the clone. Each mutation individually increased binding to the selecting H1-HA antigen, and P52aA alone conferred high affinity binding comparable to highly mutated antibodies that carry up to 22 amino-acid substitutions (Fig. 1c). Surprisingly, reversion of both mutations (R30S/A52aP) on the highly mutated antibody FI353 did not reduce binding to H1-HA. These results highlight a rapid pathway of affinity maturation through a single somatic mutation, which becomes redundant as further mutations accumulate.

To investigate the functional consequences of the intra-clonal diversification, we measured the neutralizing activity of the UCA, branchpoint and mutated antibodies on a panel of 19 group 1 influenza A viruses (Fig. 1d). While the UCA antibody failed to neutralize, the branchpoint and highly mutated antibodies showed equal capacity to neutralize the contemporary pre-pandemic H1N1 SD/07 virus. Interestingly, however, they showed different patterns of cross-neutralization of other group 1 viruses, in particular the human H2N2 JAP/57. Reversion of the two initial mutations at residues 30 and 52a on the highly mutated antibody FI353 did not affect neutralization of the contemporary virus, but reduced neutralization of other H1 strains and group 1 viruses. These results suggest that, although there is redundancy for binding to H1-HA, more mutations are required for cross-reactivity with other group 1 HAs.

Clone 1 provides another example of rapid affinity maturation followed by extensive diversification (Extended Data Fig. 3). It comprises 28 antibodies isolated over 3 years that carry from 24 to 45 somatic mutations. The first branchpoint contains six amino-acid mutations, including P52aA and S30R, which were found in the first branchpoint of clone 9. Interestingly, P52aA alone was sufficient to confer high affinity binding and neutralizing activity against H1 viruses, while P52aA and S30R together further increased binding to H1 and H5 HAs. Further mutations

¹Institute for Research in Biomedicine, Università della Svizzera Italiana, Via Vincenzo Vela 6, 6500 Bellinzona, Switzerland. ²Department of Infectious Diseases and Vaccines MedImmune LLC, One MedImmune Way, Gaithersburg, Maryland 20878, USA. ³Viral Pathogens and Biosafety Unit, San Raffaele Scientific Institute, Via Olgettina 58, 20132 Milan, Italy. ⁴Humabs BioMed SA, Via Mirasole 1, 6500 Bellinzona, Switzerland. ⁵Unit of Preventive Medicine, San Raffaele Scientific Institute, Via Olgettina 58, 20132 Milan, Italy. ⁶Institute for Microbiology, ETH Zurich, Wolfgang-Pauli-Strasse 10, 8093 Zurich, Switzerland.

*These authors contributed equally to this work.

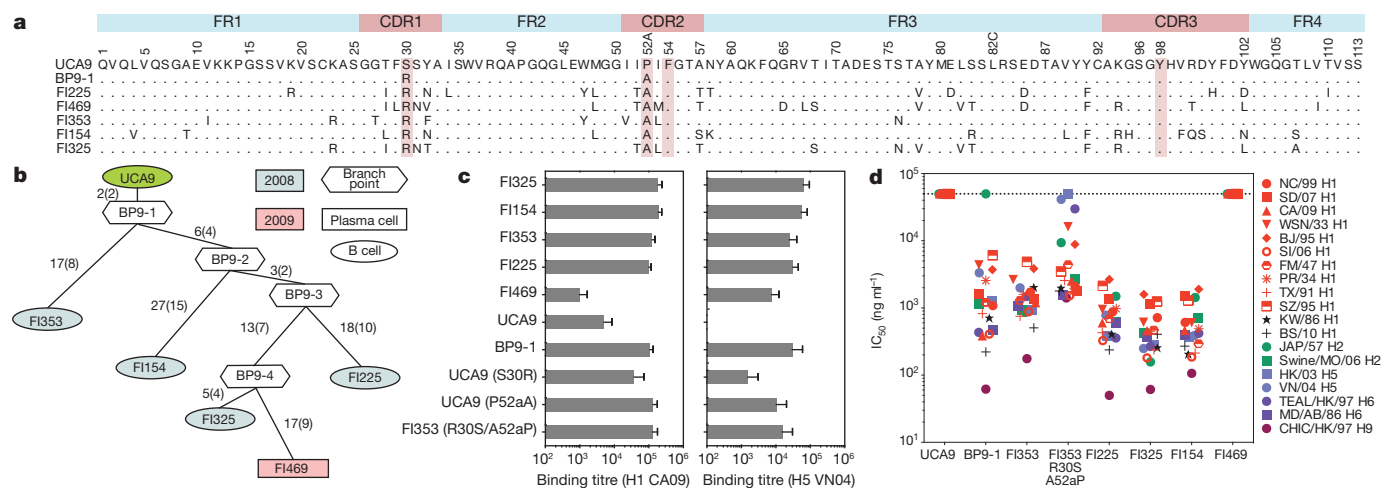


Figure 1 | Rapid affinity maturation and accumulation of redundant somatic mutations in the VH1-69 clone 9. **a**, Alignment of VH amino-acid sequences of five mutated antibodies with their UCA and branchpoint (BP) configurations. Amino-acid substitutions are highlighted in red. Residue positions are according to Kabat numbering. Dots indicate identical residues. **b**, Genealogy tree of clone 9 VH nucleotide sequences generated using dnaml. The number of mutations is indicated on the branches with amino-acid substitutions in parentheses. Background colour and shape identify the origin and the year of isolation. **c**, Binding of mutated, UCA, branchpoint and

antibody variants to H1 CA09 or H5 VN04 HAs. The mean enzyme-linked immunosorbent assay (ELISA) binding titre (EC_{50} values of a 1 mg ml⁻¹ antibody solution) of at least two independent experiments is shown. Error bars, s.e.m. Mutated residues are shown in parentheses. **d**, Neutralization of influenza A viruses. Neutralization titre (50% inhibitory concentration (IC_{50})) values above 50 μ g ml⁻¹ were scored as negative (dashed line). Complete viral strain designations are in Supplementary Fig. 6. Data represent the average of two independent experiments.

had minor effects on binding and neutralization of the selecting H1 viruses (H1N1 SD/07 and CA/09), but increased breadth of reactivity with heterologous group 1 viruses. Furthermore, reversion of the founder P52A mutation on the backbone of two highly mutated antibodies did not markedly compromise their binding to H1-HA.

While affinity maturation of clones 9 and 1 proceeds through the P52A and S30R mutations, clone 5 was found to follow a different developmental pathway. Its first branchpoint carried four amino-acid substitutions (Fig. 2a, b). Two of these mutations in the HCDR1 (T28P/S30I) were individually sufficient to confer high affinity binding to H1-HA, while a third mutation in the HCDR2 (I53V) had a moderate effect, and

a fourth (A57P) did not increase binding (Fig. 2c). Reversion of each of the two HCDR1 mutations on the mutated antibody FI3095 had minimal effects on H1-HA binding. However, reversion of both mutations on FI3095 or BP5-3 reduced binding to H1-HA. Interestingly, the reverted version of BP5-3 retained neutralizing activity towards H1 viruses, but dramatically lost reactivity against other group 1 subtypes (Fig. 2d).

The above analysis of the developmental pathways of three independent clones is consistent with a model whereby unmutated VH1-69 antibodies bind with measurable affinity to the eliciting H1-HA antigen. In all cases, high affinity binding to H1-HA was reached by a single founder mutation in either HCDR1 or HCDR2. Subsequently, the accumulation

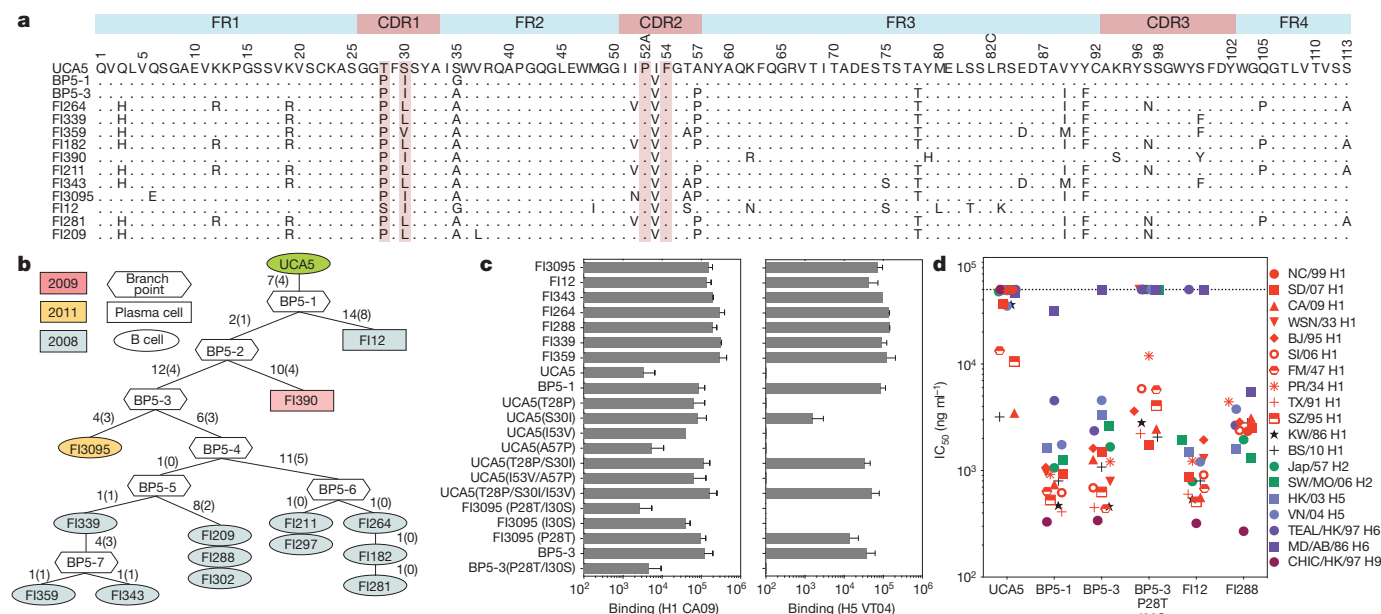


Figure 2 | Rapid affinity maturation through an alternative pathway of redundant mutations in the VH1-69 clone 5. **a**, Alignment of VH amino-acid sequences of mutated antibodies with their UCA and branchpoint configurations. **b**, Genealogy tree of clone 5 VH nucleotide sequences.

c, Binding of mutated, UCA, branchpoint and antibody variants to H1 CA09 or H5 VN04 HAs. The mean binding titre of at least two independent experiments is shown. Error bars, s.e.m. **d**, Neutralization of influenza A viruses. Data represent the average of two independent experiments.

of multiple favourable mutations diversified antibody descendants, providing alternative, redundant solutions for binding to the selecting antigen, while simultaneously conferring an overall broader reactivity against unrelated group 1 HAs.

To identify conserved residues and constrained mutations that may have a role in antibody binding, we analysed the amino-acid usage in the HCDR1 and HCDR2 in a panel of 119 anti-stem VH1-69 antibodies isolated in our laboratory or described in the literature (Supplementary Fig. 7). A previous study suggested that the flipping out of F29, induced by HFR3 mutations, was required for HA binding⁹. However, we found that F29 is not conserved in all antibodies analysed, and that F29A substitutions on several antibodies did not affect binding to H1-HA (Extended Data Fig. 4). In contrast, we found that some residues (G26, A33, I51, G55 and F54) show more than 90% conservation and a low replacement to silent mutation ratio (R/S), consistent with negative antigenic selection (Fig. 3a, b and Extended Data Fig. 6). Moreover, in most antibodies I53 is mutated to other hydrophobic amino acids and A57 to threonine or proline. P52a was exclusively mutated to alanine or glycine. Interestingly, in the three crystal structures solved¹⁻³, the residues at positions 52a and 57 are not in contact with the HA antigen, suggesting that mutations in these positions could have an allosteric effect. In contrast, in the three structures F54 is a conserved contact site that occupies the W21 aromatic pocket. The role of F54 is also evident from its absolute conservation in the panel of VH1-69 antibodies analysed. Of note, position

54 is polymorphic, with 6 out of the 14 VH1-69 alleles encoding for leucine instead of phenylalanine^{13,14} (Fig. 3c and Extended Data Fig. 7). Consistently, all anti-stem antibodies described so far derive from the F54-encoding alleles *01, *03, *06 and *12 (Supplementary Figs 1 and 8).

To examine the impact of position 54 polymorphism on the antibody response to the HA stem, we genotyped 345 volunteers and measured the level of serum antibodies that inhibit binding of an anti-stem VH1-69 antibody to H1-HA⁷. As recently reported¹⁵, the titre of anti-stem antibodies was age-dependent, with individuals older than 40 years showing significantly higher levels (Fig. 3d). Eleven per cent of the donors were homozygous for L54 alleles (L/L) while the remaining were either heterozygous (56%, F/L) or homozygous for F54 alleles (33%, F/F). Interestingly, in younger individuals (<40 years), the levels of serum anti-stem antibodies were significantly lower in the L/L compared with the F/F and F/L genotype groups (Fig. 3e). These data are consistent with the major, albeit not exclusive, contribution of F54-encoding VH1-69 alleles to the development of anti-stem antibodies. This interpretation is supported by the finding that VH1-69 is not the only VH gene used to generate anti-stem antibodies⁷ (Supplementary Fig. 1), and by the exclusive isolation of non-VH1-69 anti-stem antibodies from an L/L donor (B.F.-R., unpublished observations).

The conservation of F54 in all VH1-69 anti-stem antibodies, together with the genetic and structural data, suggest that this residue plays an essential role in the development of anti-stem antibodies. Having

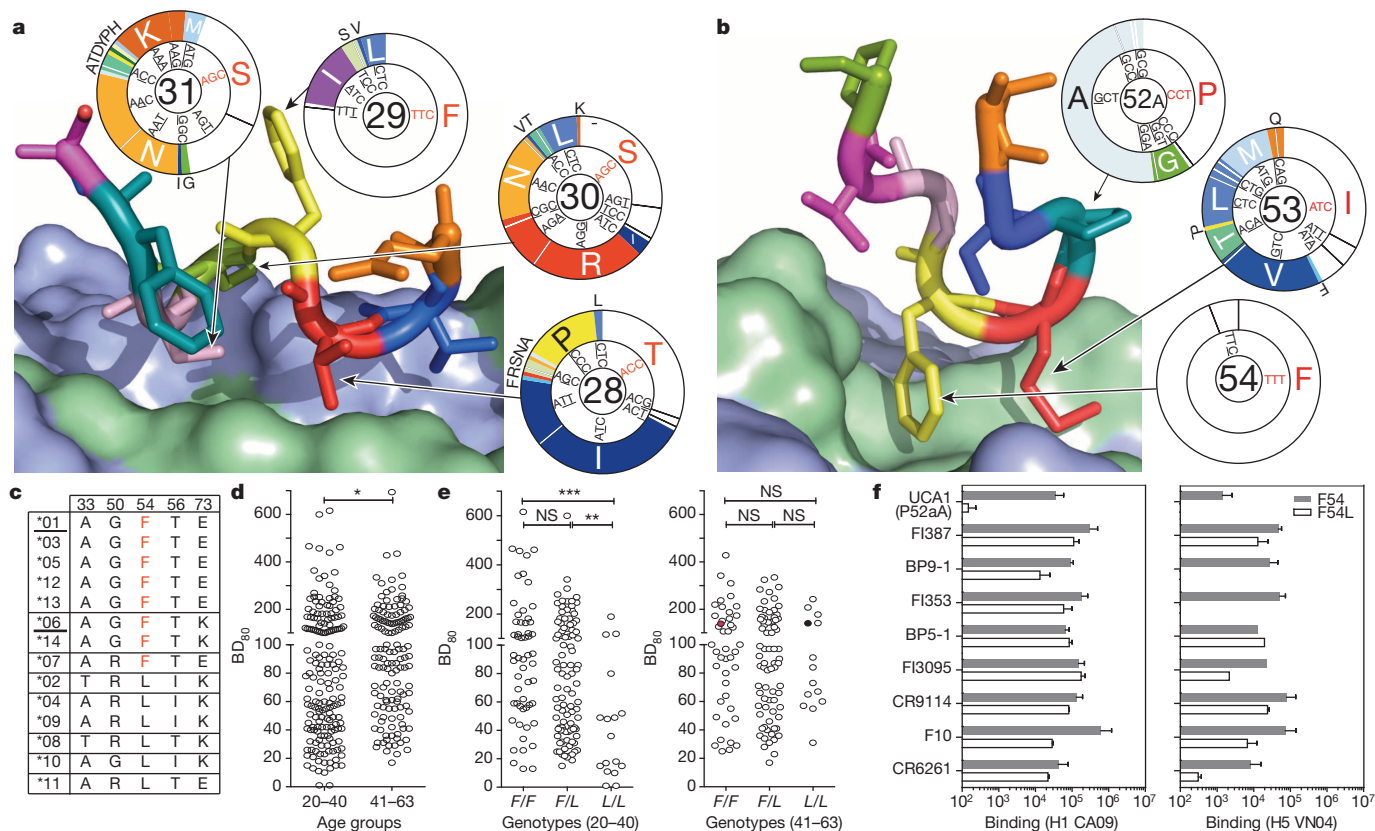


Figure 3 | Genetic requirements and mutational constraints in the HCDR1 and HCDR2 of VH1-69 anti-stem antibodies and the redundant role of F54.

a, b, Model of HCDR1 and HCDR2 of the F10 antibody bound to the H5 VN04 (Protein Data Bank (PDB) accession number 3FKU). Pie charts indicate the frequency of UCA (white) and mutated residues (International ImMunoGeneTics Information System (IMGT) colours), as well as codon usage. UCA codons and amino acids are in red. Fully annotated pie charts for all HCDR1 and HCDR2 residues are shown in Extended Data Fig. 5. **c**, Polymorphic amino-acid residues encoded by the 14 VH1-69 alleles (Extended Data Fig. 7). Alleles *01 and *06 are present as duplicated genes on the same chromosome^{13,14}. **d**, Titres of serum antibodies that inhibit the binding

of the labelled anti-stem antibody FE43. The graphs depict the serum titre that causes 80% inhibition of FE43 binding (BD_{80}) in two age groups (20–40 and >40 years). The two-tailed P value calculated with the unpaired Mann–Whitney U -test ($*P \leq 0.05$) is shown. **e**, BD_{80} values in the two age groups divided by genotype. The red dot indicates the donor analysed in this study. The black dot indicates a donor from whom only non-VH1-69 anti-stem antibodies were isolated. The P value of a Kruskal–Wallis test ($***P \leq 0.001$; $**P \leq 0.01$; $*P \leq 0.05$; NS, $P > 0.05$) is shown. **f**, Binding of mutated, branchpoint and their F54L variant antibodies to H1 CA09 or H5 VN04 HAs. The mean titre of at least two independent experiments is shown. Error bars, s.e.m.

discovered that the founder mutations become redundant in the context of fully mutated antibodies, we tested whether this could also apply to the germline-encoded F54. We therefore replaced phenylalanine 54 with either leucine or alanine in branchpoint and mutated antibodies representative of clones 1, 5 and 9, and in the three antibodies for which a crystal structure has been solved (F10, CR6261 and CR9114). Strikingly, in all cases we observed that a F54L mutation did not result in a substantial loss of binding to the selecting H1-HA, although binding to heterologous H5-HA was affected (Fig. 3f and Extended Data Fig. 8). These findings indicate that F54 is essential for initial recognition of HA by anti-stem antibodies, but can become redundant as favourable mutations accumulate.

Analysis of the antibody panel (Supplementary Fig. 9) showed that most anti-stem VH1-69 antibodies are characterized by short HCDRs,

typically 13–14 amino acids long, with Y at position 98 in over 60% of cases. In contrast, VH1-69 antibodies of different specificities showed a broader distribution of HCDR3 lengths and were not enriched for aromatic residues at position 98 (Fig. 4a, b). Structural alignment of F10, CR6261 and CR9114 revealed an identical positioning of Y98 in the W21 aromatic pocket, in spite of differences in the length and structure of their HCDR3s (Fig. 4c). Y98A substitutions in F10, CR6261, CR9114 and in the branchpoints of clones 1 and 9 reduced binding to H1-HA and H5-HA, while Y98F did not affect binding of BP9-1, indicating that the main contribution for binding is due to aromatic interactions (Fig. 4d). Taken together, these results suggest that Y98 represents an important motif for most VH1-69 antibodies, as recently described¹⁶. Interestingly, Y98-bearing clones 1 and 9 show a convergent evolutionary pathway, through the common acquisition of S30R/P52aA mutations in the HCDR1 and HCDR2. In contrast, clones lacking Y98, such as clone 5, adopt different solutions since they lack the P52aA mutation and have other characteristic mutations in the HCDR1 (Extended Data Fig. 10a).

Our study identifies three main requirements for the development of broadly influenza neutralizing antibodies: (1) a VH1-69 allele encoding F54, (2) a permissive HCDR3, which in most cases carries Y98, and (3) a single somatic mutation in the HCDR1 or HCDR2 that increases the affinity to maximal levels. The major developmental pathway, exemplified by clones 9 and 1 (as well as 2, 14 and 19), involves a P52aA mutation on Y98-bearing antibodies. This founder mutation is not a direct contact residue and therefore may influence either the overall flexibility of HCDRs or frameworks, as suggested for antibodies to HIV-1 and the H1-HA globular head^{17,18}, or the thermodynamic stability of the antibody⁸. These hypotheses could be addressed by long-timescale molecular dynamics calculations. Previous studies defined a threshold and a ceiling to affinity maturation based on the dissociation rate of the B-cell receptor antigen complex^{19–21}. According to our SPR analysis, unmutated VH1-69 antibodies have low affinity, but dissociation rates above the estimated threshold for antigen presentation, while the branchpoint antibodies have already reached the estimated ceiling characteristic of highly mutated antibodies.

The rapid development of VH1-69 anti-stem antibodies through a single somatic mutation contrasts with the requirement for a large number of somatic mutations in HIV-1 broadly neutralizing antibodies²² and has important implications for the development of a universal influenza vaccine. Once generated, the anti-stem B cells may further expand and diversify in response to repeated antigenic stimulation by re-entering into germinal centres^{23–26}. This model is consistent with the finding that anti-stem antibodies are found at higher titres in older individuals and carry a higher load of somatic mutations compared with antibodies specific for variable epitopes of the globular head.

Somatic mutations can broaden antibody reactivity against HIV-1, influenza and paramyxoviruses^{4,12,27}. However, while antibody breadth to HIV-1 and paramyxoviruses might result from serial exposure to variant antigens, in the case of influenza the cross-reactivity to animal HA subtypes, such as H5, H6 or H9, is unlikely to result from exposure to these antigens. Our analysis suggests a stepwise process of diversification that starts with the rapid acquisition of high affinity binding to the eliciting H1-HA, followed by the accumulation of additional mutations conferring breadth to different HA subtypes (Extended Data Fig. 10b). We speculate that the extensive intraclonal diversification generated by continuous somatic mutations may provide a repertoire of antibodies with slightly different paratopes potentially able to cope with rapidly diversifying viruses.

An interesting and unexpected finding of our study is that the founder mutations that maximize binding to the eliciting antigen, and even the germline-encoded F54 that is required to initiate affinity maturation, become redundant through the accumulation of additional favourable mutations. We suggest that this mechanism provides a robust solution to maintain high affinity binding while allowing repertoire diversification, and represents a new way of understanding the process of affinity maturation in the germinal centre reaction.

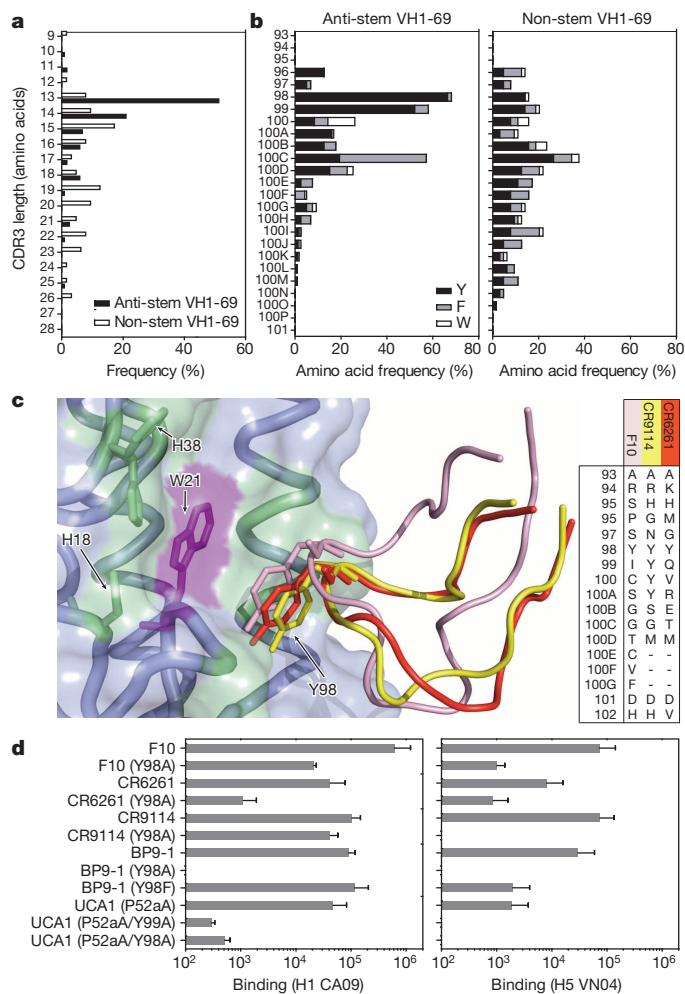


Figure 4 | Constraints in the HCDR3 of VH1-69 anti-stem antibodies.

a, HCDR3 length distribution in VH1-69 antibodies specific for the HA stem ($n = 119$) or for unrelated antigens ($n = 64$). The complete list of non-HA specific antibodies used is shown in Supplementary Fig. 10. **b**, Frequency of aromatic residues in the HCDR3 of VH1-69 antibodies specific for the HA stem (left panel) or for unrelated antigens (right panel). **c**, Superimposition of the HCDR3 of F10 (pink, PDB 3FKU), CR6261 (red, PDB 3GBM) and CR9114 (yellow, PDB 4FQI) relative to H5 VN04 HA derived by the structural alignment of HA2 atoms on PyMOL. The antibody Y98 residues and the conserved HA aromatic residues H18, H38 and W21 are shown as sticks. HCDR3 amino-acid sequences are also shown. HCDR2 amino-acid structural alignment is shown in Extended Data Fig. 9. **d**, Binding of mutated, branchpoint and their HCDR3 variant antibodies to H1 CA09 or H5 VN04 HAs. The mean titre of at least two independent experiments is shown. Error bars, s.e.m.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 28 April; accepted 11 August 2014.

Published online 5 October 2014.

- Sui, J. *et al.* Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. *Nature Struct. Mol. Biol.* **16**, 265–273 (2009).
- Ekiert, D. C. *et al.* Antibody recognition of a highly conserved influenza virus epitope. *Science* **324**, 246–251 (2009).
- Dreyfus, C. *et al.* Highly conserved protective epitopes on influenza B viruses. *Science* **337**, 1343–1348 (2012).
- Corti, D. *et al.* A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. *Science* **333**, 850–856 (2011).
- Throsby, M. *et al.* Heterosubtypic neutralizing monoclonal antibodies cross-protective against H5N1 and H1N1 recovered from human IgM+ memory B cells. *PLoS ONE* **3**, e3942 (2008).
- Wrammert, J. *et al.* Broadly cross-reactive antibodies dominate the human B cell response against 2009 pandemic H1N1 influenza virus infection. *J. Exp. Med.* **208**, 181–193 (2011).
- Corti, D. *et al.* Heterosubtypic neutralizing antibodies are produced by individuals immunized with a seasonal influenza vaccine. *J. Clin. Invest.* **120**, 1663–1673 (2010).
- Wang, F. *et al.* Somatic hypermutation maintains antibody thermodynamic stability during affinity maturation. *Proc. Natl Acad. Sci. USA* **110**, 4261–4266 (2013).
- Lingwood, D. *et al.* Structural and genetic basis for development of broadly neutralizing influenza antibodies. *Nature* **489**, 566–570 (2012).
- Jacob, J., Kelsoe, G., Rajewsky, K. & Weiss, U. Intracloonal generation of antibody mutants in germinal centres. *Nature* **354**, 389–392 (1991).
- McKean, D. *et al.* Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. *Proc. Natl Acad. Sci. USA* **81**, 3180–3184 (1984).
- Liao, H.-X. *et al.* Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* **496**, 469–476 (2013).
- Sasso, E. H., Johnson, T. & Kipps, T. J. Expression of the immunoglobulin VH gene 51p1 is proportional to its germline gene copy number. *J. Clin. Invest.* **97**, 2074–2080 (1996).
- Vencovský, J. *et al.* Polymorphism in the immunoglobulin VH gene V1-69 affects susceptibility to rheumatoid arthritis in subjects lacking the HLA-DRB1 shared epitope. *Rheumatology* **41**, 401–410 (2002).
- Sangster, M. Y. *et al.* B cell response and hemagglutinin stalk-reactive antibody production in different age cohorts following 2009 H1N1 influenza virus vaccination. *Clin. Vaccine Immunol.* **20**, 867–876 (2013).
- Avnir, Y. *et al.* Molecular signatures of hemagglutinin stem-directed heterosubtypic human neutralizing antibodies against influenza A viruses. *PLoS Pathog.* **10**, e1004103 (2014).
- Klein, F. *et al.* Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. *Cell* **153**, 126–138 (2013).
- Schmidt, A. G. *et al.* Preconfiguration of the antigen-binding site during affinity maturation of a broadly neutralizing influenza virus antibody. *Proc. Natl Acad. Sci. USA* **110**, 264–269 (2013).
- Batista, F. & Neuberger, M. Affinity dependence of the B cell response to antigen: a threshold, a ceiling, and the importance of off-rate. *Immunity* **8**, 751–759 (1998).
- Foote, J. & Eisen, H. N. Kinetic and affinity limits on antibodies produced during immune responses. *Proc. Natl Acad. Sci. USA* **92**, 1254–1256 (1995).
- Poulsen, T. R., Jensen, A., Haurum, J. S. & Andersen, P. S. Limits for antibody affinity maturation and repertoire diversification in hypervaccinated humans. *J. Immunol.* **187**, 4229–4235 (2011).
- Kwong, P. D., Mascola, J. R. & Nabel, G. J. Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning. *Nature Rev. Immunol.* **13**, 693–701 (2013).
- Victoria, G. D. & Nussenzweig, M. C. Germinal centers. *Annu. Rev. Immunol.* **30**, 429–457 (2012).
- Dogan, I. *et al.* Multiple layers of B cell memory with different effector functions. *Nature Immunol.* **10**, 1292–1299 (2009).
- Pape, K. A., Taylor, J. J., Maul, R. W., Gearhart, P. J. & Jenkins, M. K. Different B cell populations mediate early and late memory during an endogenous immune response. *Science* **331**, 1203–1207 (2011).
- Kaji, T. *et al.* Both mutated and unmutated memory B cells accumulate mutations in the course of the secondary response and develop a new antibody repertoire optimally adapted to the secondary stimulus. *Int. Immunol.* **25**, 683–695 (2013).
- Corti, D. *et al.* Cross-neutralization of four paramyxoviruses by a human monoclonal antibody. *Nature* **501**, 439–443 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was partly supported by the European Research Council (grant number 250348) IMMUNExplore, the Swiss National Science Foundation (grant number 141254), Fondazione Cariplo Vaccine Program (grant number 2009-3594), the Human Frontiers Science Program (grant number RGP0009/2007-C), the European Commission (grant number FP7-HEALTH-2011-280873) ADITEC and the National Institutes of Health (U19 grant number AI-057266). A.L. is supported by the Helmut Horten Foundation.

Author Contributions L.Pa. designed and performed experiments, analysed the data and wrote the manuscript. L.Pa. and M.F. performed phylogenetic analysis. M.F. performed bioinformatic analysis. L.Pi. performed SPR and IF experiments. C.S., B.F.-R., G.A. and I.G.-S. provided technical support for antibody isolation and purification. F.T. and E.V. performed genotypic analysis. N.L.K. and Q.Z. designed, performed and analysed viral neutralization experiments. G.P. provided blood samples from healthy volunteers. F.S. wrote the manuscript. D.C. and A.L. designed the experiments, provided overall supervision and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.L. (lanzavecchia@irb.usi.ch).

METHODS

Viruses and antigens. Wild-type influenza strains were obtained from the Centers for Disease Control and Prevention or purchased from the American Tissue Culture Collection. Cold adapted (*ca*) live-attenuated influenza vaccine viruses were generated by either classical reassortment or by reverse genetics²⁸. All viruses were propagated in embryonated chicken eggs, and virus titres were determined by median tissue culture infective dose (TCID₅₀) per millilitre. The viruses used in this study were H1N1 A/Caledonia/20/99, H1N1 A/South Dakota/6/2007, H1N1 A/California/7/2009, H1N1 A/Wilson Smith N/33, H1N1 *ca* A/Beijing/262/95, H1N1 A/Solomon Islands/3/2006, H1N1 A/Fort Monmouth/1/47, H1N1 A/Puerto Rico/8/34, H1N1 *ca* A/Texas/36/91, H1N1 *ca* A/Shenzhen/227/95, H1N1 *ca* A/Kawasaki/9/86, H1N1 A/Brisbane/10/2010, H2N2 *ca* A/Japan/57, H2N3 *ca* A/swine/Missouri/4296424/2006, H5N1 *ca* A/Hong Kong/213/2003, H5N1 *ca* A/Vietnam/1203/2004, H6N1 *ca* A/Hong Kong/W312/97, H6N2 *ca* A/mallard/Alberta/89/85, H9N2 *ca* A/chicken/Hong Kong/G9/97. Replication-incompetent virus pseudotyped with the HA genes of H5N1 A/Vietnam/1203/04 were produced by co-transfection of HEK293T/17 with an H5 HA-expressing plasmid and a complementing viral-genome reporter vector, pNL4-3.Luc⁺.E.R⁺ (provided by J. R. Mascola) in the presence of 0.1 U ml⁻¹ recombinant neuraminidase from *Clostridium perfringens* (Sigma) as described previously²⁹. The antigens used in the ELISA and SPR experiments were A/H1/California/07/2009 (Protein Sciences), A/H5/Vietnam/1203/2004 (Protein Sciences) and Tetanus Toxoid (Sigma).

Sample collection. Peripheral blood mononuclear cell (PBMC) samples were obtained from a single healthy donor after vaccination in December 2008 (2008–2009 season vaccine), January 2010 (2009–2010 season vaccine), December 2010 (2010–2011 season vaccine), November–December 2011 (2011–2012 season vaccine) and December 2012 (2012–2013 season vaccine), as well as following infection with swine-origin influenza virus in November 2009. Memory B cells were isolated from fresh or cryopreserved PBMCs. Plasma cells were isolated from fresh PBMCs. The donor gave written informed consent for the use of these blood samples, following approval by the Cantonal Ethical Committee of Canton Ticino, Switzerland. Following written informed consent signature, a blood sample donation was obtained from 345 workers of the San Raffaele Hospital and Scientific Institute, Milan, Italy. Both genders and individuals between 20 and 63 years of age were included in the study. Serum was separated from whole venous blood and stored at –80 °C until use. DNA samples were obtained from either blood clots or PBMCs as previously described³⁰. All procedures were approved by the review board of the San Raffaele Hospital and Scientific Institute.

Monoclonal antibody isolation from memory B cells. Memory B cells were isolated from cryopreserved or fresh PBMCs using CD22 microbeads or anti-FITC (fluorescein isothiocyanate) microbeads (Miltenyi Biotec) after staining of PBMCs with CD22-FITC, and were immortalized with Epstein–Barr virus (EBV) and CpG in multiple wells as described previously³¹. Plasma cells from peripheral blood were stained with anti-CD138 antibody conjugated to phycoerythrin (PE) (BD-Pharmingen), enriched by magnetic separation with anti-PE microbeads (Miltenyi) and purified by cell sorting on a FACSAria (BD Biosciences). Plasma cells were seeded at 0.5 cells per well as previously described⁴. Supernatants of plasma cells cultured for 3–4 days and supernatants of immortalized memory B cells collected after 2 weeks were screened for their ability to bind H5 HA or neutralize H5 pseudovirus infections in high-throughput ELISA and micro-neutralization assays, respectively. Positive EBV-B cell cultures were expanded in complete RPMI medium. VH and VL sequences were obtained from positive plasma cell and B-cell cultures by reverse transcriptase PCR (RT–PCR). Selected sequences were cloned into human immunoglobulin-G1 (IgG1) and Igk or Igλ expression vectors as described previously³². Vectors were provided by M. Nussenzweig. Monoclonal antibodies were produced by transient transfection of 293F cells (Invitrogen) using PEI in serum-free media. Cell supernatants were collected 4–8 days after transfection, filtered and preserved by the addition of 0.1% BSA/PBS. The antibodies were affinity purified by protein A chromatography (GE Healthcare).

Virus neutralization. The microneutralization assay was modified from a previously described accelerated viral inhibition assay with neuraminidase as read-out³³. Briefly, 100 TCID₅₀ of virus was added to threefold dilutions of antibody in a 384-black walled plate and incubated for 1 h incubation at 33 °C. After incubation, 2 × 10⁴ Madin Darby canine kidney cells per well were added to the plate, then further incubated at 33 °C in a CO₂ incubator for approximately 40 h. Neuraminidase activity was measured by adding a fluorescently labelled substrate, methylumbelliferyl-N-acetyl neuraminic acid to each well and incubated at 37 °C for 1 h. Virus replication represented by neuraminidase activity was quantified by reading fluorescence with an Envision Fluorometer (PerkinElmer) using the following settings: excitation 355 nm; emission 460 nm; ten flashes per well. The neutralization titre (50% inhibitory concentration (IC₅₀)) is expressed as the antibody concentration that reduced the fluorescence signal by 50% compared with cell control wells.

ELISA. ELISA high protein binding plates (Perkin Elmer) were coated with 6 µg ml⁻¹ of recombinant HA proteins (Protein Sciences), and Tetanus Toxoid (Sigma) or PBS (Gibco) as a control, and incubated overnight at 4 °C. Plates were blocked using a 1% w/v solution of bovine serum albumin (Sigma) in PBS for 1 h at 20–25 °C. Serial dilutions of the antibodies were incubated for 1.5 h at room temperature. After washing, antibody binding was revealed using a secondary F(ab)₂ goat anti-human IgG antibody conjugated to alkaline peroxidase (Southern Biotechnologies). Plates were then washed, substrate (p-NPP, Sigma) was added and plates were read at 405 nm. The relative affinities of antibody binding were determined by measuring the concentration of antibody required to achieve 50% binding relative to the maximum (EC₅₀). The EC₅₀ values were calculated by interpolation of binding curves fitted with a four-parameter nonlinear regression with a variable slope. All antibodies were tested in duplicate.

SPR. The kinetic parameters of the binding of antibodies to H1-HA were determined by SPR. Protein A (450 nM) was stabilized in 10 mM acetate buffer, pH 4.5, and immobilized onto an ethyl(dimethylaminopropyl) carbodiimide/N-Hydroxysuccinimide (EDC/NHS) pre-activated ProteOn sensor chip (Bio-Rad) through amine coupling; unreacted groups were blocked by injection of 1 M ethanolamine HCl. HEPES-buffered saline (HBS) (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% surfactant Tween-20) was used as a running buffer. All injections were made at a flow rate of 100 µl min⁻¹. Monoclonal antibodies were diluted in HBS to 10 nM and injected for 30 s onto the protein-A-coated chip for capturing, followed by injection of different concentrations (30 nM, 10 nM, 3.3 nM, 1.1 nM, 0.37 nM) of A/H1/California/07/09 (Protein Sciences). Injection time was 240 s and dissociation time was 900 s. Antibodies that did not show clear binding were re-tested using more sensitive conditions: antibodies (100 nM) injected for 180 s to saturate the protein A chip, followed by injection of higher concentrations of A/H1/California/07/09 (300 nM, 100 nM, 33.3 nM, 11.1 nM, 3.7 nM). Injection time was 240 s and dissociation time was 600 s. Regeneration of the sensor chip was done by double pulses of H₃PO₄ (10 mM). One channel of the chip was injected with HBS and used as reference for the analysis. Each binding interaction of the monoclonal antibodies to H1-HA was assessed using a ProteOn XPR36 instrument (Bio-Rad) and the data were processed with ProteOn Manager software. The acid dissociation constant (K_a), dissociation constant (K_d) and equilibrium dissociation constant (K_D) values were calculated applying the Langmuir fit model.

Genotype analysis of F54L. The VH1-69 gene was amplified from DNA extracted from blood samples of 345 individuals by PCR using the primer pair A, 5'-GAGG AAGGATCCTGGTT-3', and B, 5'-GGATGTGGGTTTTCACACTGTG-3', as previously described⁹. Direct sequencing of the amplified DNA was performed using primers A and B in an ABI PRISM 3730 DNA Analyzer. The electropherograms of each sequence were analysed for allelic discrimination. As a control, 34 samples were randomly selected, re-amplified and sequenced. No significant discrepancy was observed between the results of the two independent analyses.

Immunoglobulin lineage and sequence analysis. The DNA Maximum Likelihood program (Dnaml) of the PHYLIP package, version 3.69, was used to estimate immunoglobulin phylogenies from nucleotide sequences that were first aligned using ClustalW2 (ref. 34), as previously described¹². IgH CDR3 regions were defined with their Kabat numbering (positions 93–102) using the software available on the Abnum website (<http://www.bioinf.org.uk/abs/abnum/>)³⁵. IgH CDR1 and CDR2 sequences were aligned using ClustalW2, and IgH CDR3 sequences were listed according to the alignment of the immunoglobulin CDR1 and CDR2. The alignments were displayed in a graphic user interface written in Java. The V, D and J genes of the IgH DNA sequences were identified using IgBlast and the IMGT database as a reference^{36,37}.

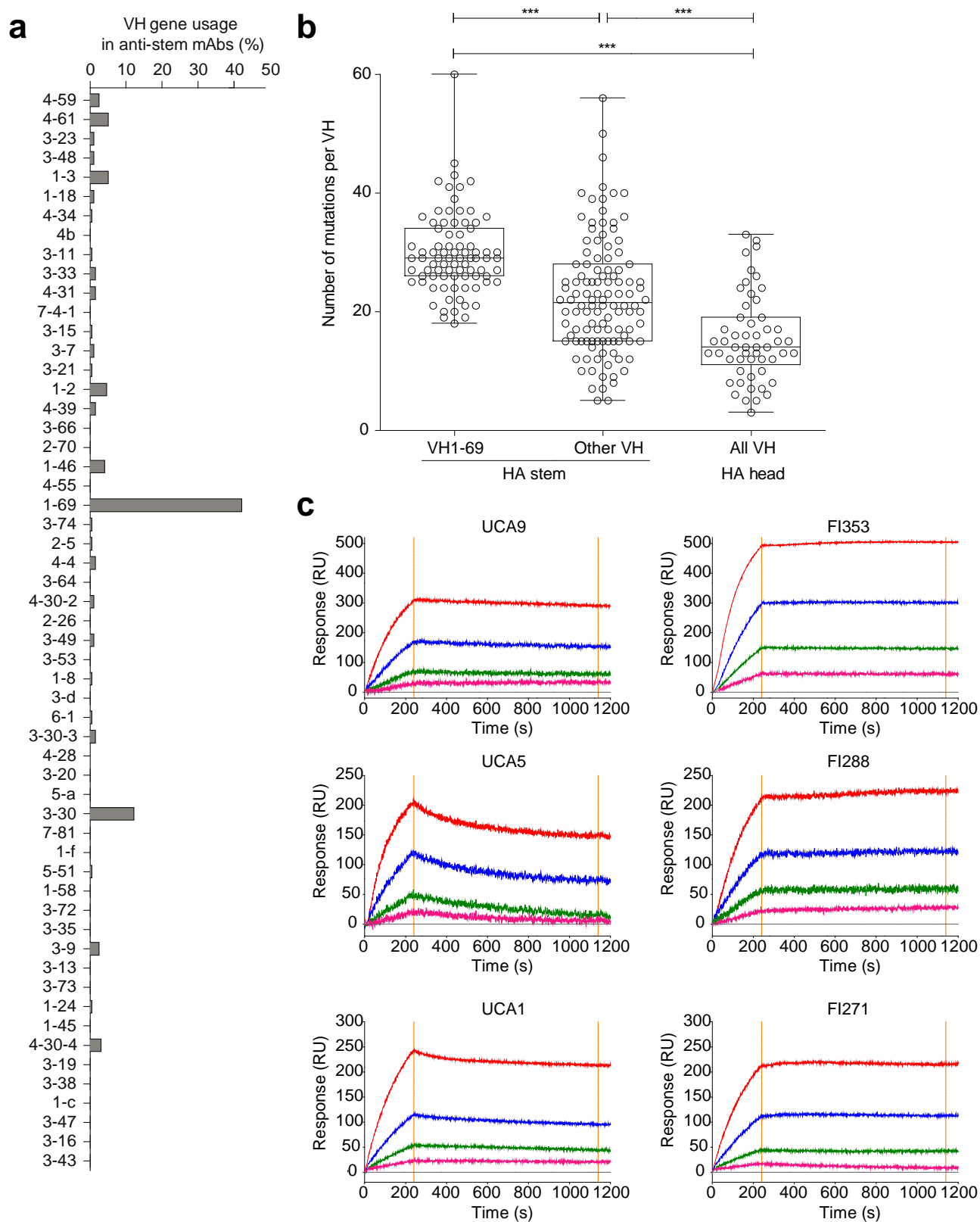
Structural analysis. Structural alignments of the antibodies CR9114, CR6261 and F10 in complex with H5 VN04 were obtained by aligning the HA2 of H5 VN04 of the structures deposited in PDB under accession numbers 4FQI³, 3GBM² and 3FKU¹, respectively. These structures were determined by X-ray crystallography at resolutions of 1.71, 2.7 and 3.2 Å, respectively. For visualization purposes, only the trace HCDR3 backbone of the three antibodies is depicted.

Antibody variants. UCA sequences were determined with reference to the IMGT database, and produced by gene synthesis (Genscript). The HCDR3 of the UCA was defined as the IMGT-derived rearrangement that required a minimal number of HCDR3 mutations to be introduced downstream in the tree. For certain antibodies, alternative unmutated versions were synthesized using the HCDR3 of the wild-type antibody. Branchpoint antibodies for multi-member clones were determined by the antibody phylogenetic tree derived by dnaml, and produced by gene synthesis (Genscript) or by site-directed mutagenesis (Agilent, Quick Change Lightning Site Directed Mutagenesis kit). CR9114 and CR6261 V_H and V_L sequences were obtained through GenBank (accession numbers JX213639.1, JX213640.1, HI919029.1, HI919031.1) and F10 V_H and V_L sequences were obtained through PDB (accession number 3FKU).

Statistics. The Mann–Whitney two-tailed unpaired U-test was used for statistical comparisons of two data groups, and the Kruskal–Wallis test was used for statistical

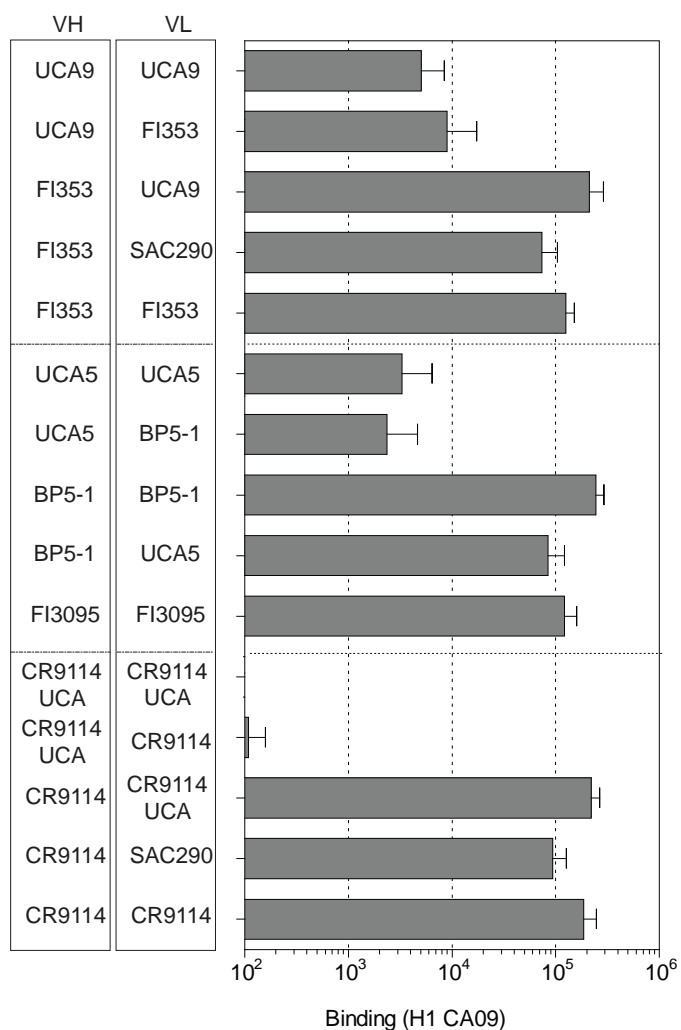
comparisons of multiple data groups; *P* values of 0.05 or less were considered significant.

28. Jin, H. *et al.* Multiple amino acid residues confer temperature sensitivity to human influenza virus vaccine strains (FluMist) derived from cold-adapted A/Ann Arbor/6/60. *Virology* **306**, 18–24 (2003).
29. Temperton, N. J. *et al.* A sensitive retroviral pseudotype assay for influenza H5N1-neutralizing antibodies. *Influenza Other Respi. Viruses* **1**, 105–112 (2007).
30. Ghezzi, S. *et al.* Identification of TRIM22 single nucleotide polymorphisms associated with loss of inhibition of HIV-1 transcription and advanced HIV-1 disease. *AIDS* **27**, 2335–2344 (2013).
31. Traggiai, E. *et al.* An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. *Nature Med.* **10**, 871–875 (2004).
32. Tiller, T. *et al.* Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J. Immunol. Methods* **329**, 112–124 (2008).
33. Hassantoufighi, A. *et al.* A practical influenza neutralization assay to simultaneously quantify hemagglutinin and neuraminidase-inhibiting antibody responses. *Vaccine* **28**, 790–797 (2010).
34. Goujon, M. *et al.* A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* **38**, W695–W699 (2010).
35. Abhinandan, K. R. & Martin, A. C. R. Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains. *Mol. Immunol.* **45**, 3832–3839 (2008).
36. Ye, J., Ma, N., Madden, T. L. & Ostell, J. M. IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res.* **41**, W34–W40 (2013).
37. Brochet, X., Lefranc, M.-P. & Giudicelli, V. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res.* **36**, W503–W508 (2008).
38. Arnaout, R. *et al.* High-resolution description of antibody heavy-chain repertoires in humans. *PLoS ONE* **6**, e22365 (2011).
39. Sasso, E. H., Willems van Dijk, K., Bull, A. P. & Milner, E. C. A fetally expressed immunoglobulin VH1 gene belongs to a complex set of alleles. *J. Clin. Invest.* **91**, 2358–2367 (1993).

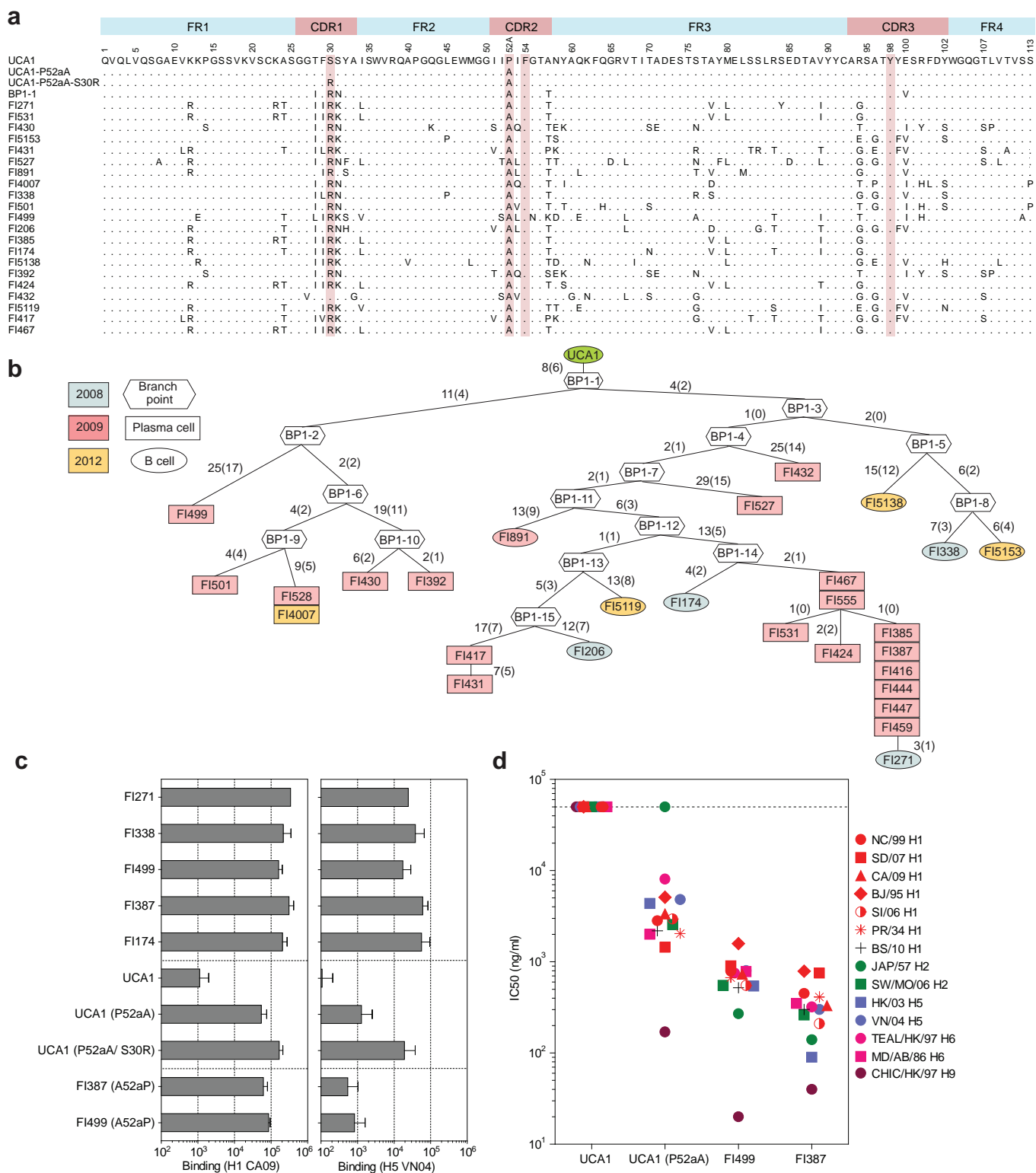


Extended Data Figure 1 | VH1-69 antibodies dominate the response to the HA stem and are highly mutated. **a**, Frequency of VH gene usage in 197 anti-stem antibodies isolated from the donor analysed in this study. The VH genes are listed in descending order according to their frequency of usage in normal adult human PBMCs as described in ref. 38. **b**, Load of somatic mutations in the VH of anti-stem antibodies using VH1-69 or other VH genes,

compared with anti-head antibodies specific for the H1 CA09 HA. Two-tailed P value was calculated with an unpaired Student's t -test. *** $P \leq 0.001$; * $P \leq 0.05$. **c**, Binding of the UCA IgG antibody (left) and the corresponding mutated antibody (right) to H1 CA09 HA from three representative clones as measured by SPR (complete data set in Supplementary Figs 2 and 3). RU, resonance units.

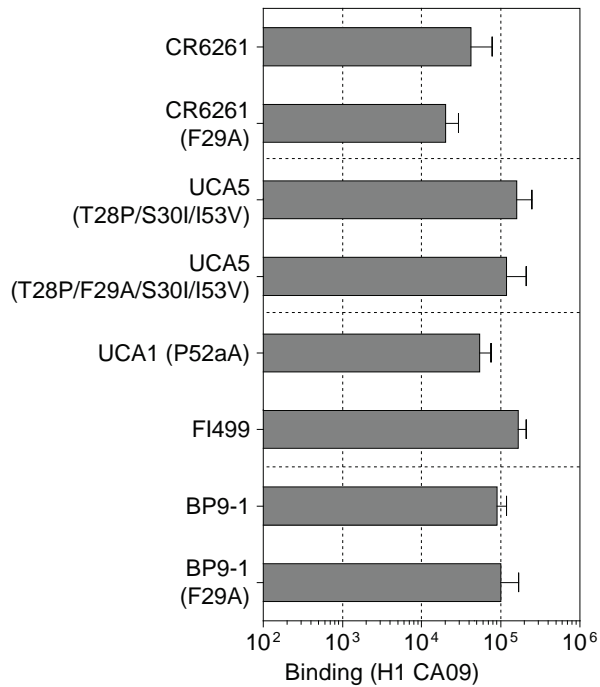


Extended Data Figure 2 | Light chains do not contribute to the binding of VH1-69 antibodies to the HA stem. Recombinant antibodies were produced using different combinations of VH and VL including a light chain from an antibody of a different specificity, SAC290. The binding of mutated, branchpoint, UCA and VH/VL shuffled antibodies to H1 CA09 HA as measured by ELISA (mean binding titre value of a 1 mg ml⁻¹ antibody solution \pm s.e.m. in at least two independent experiments) is shown.

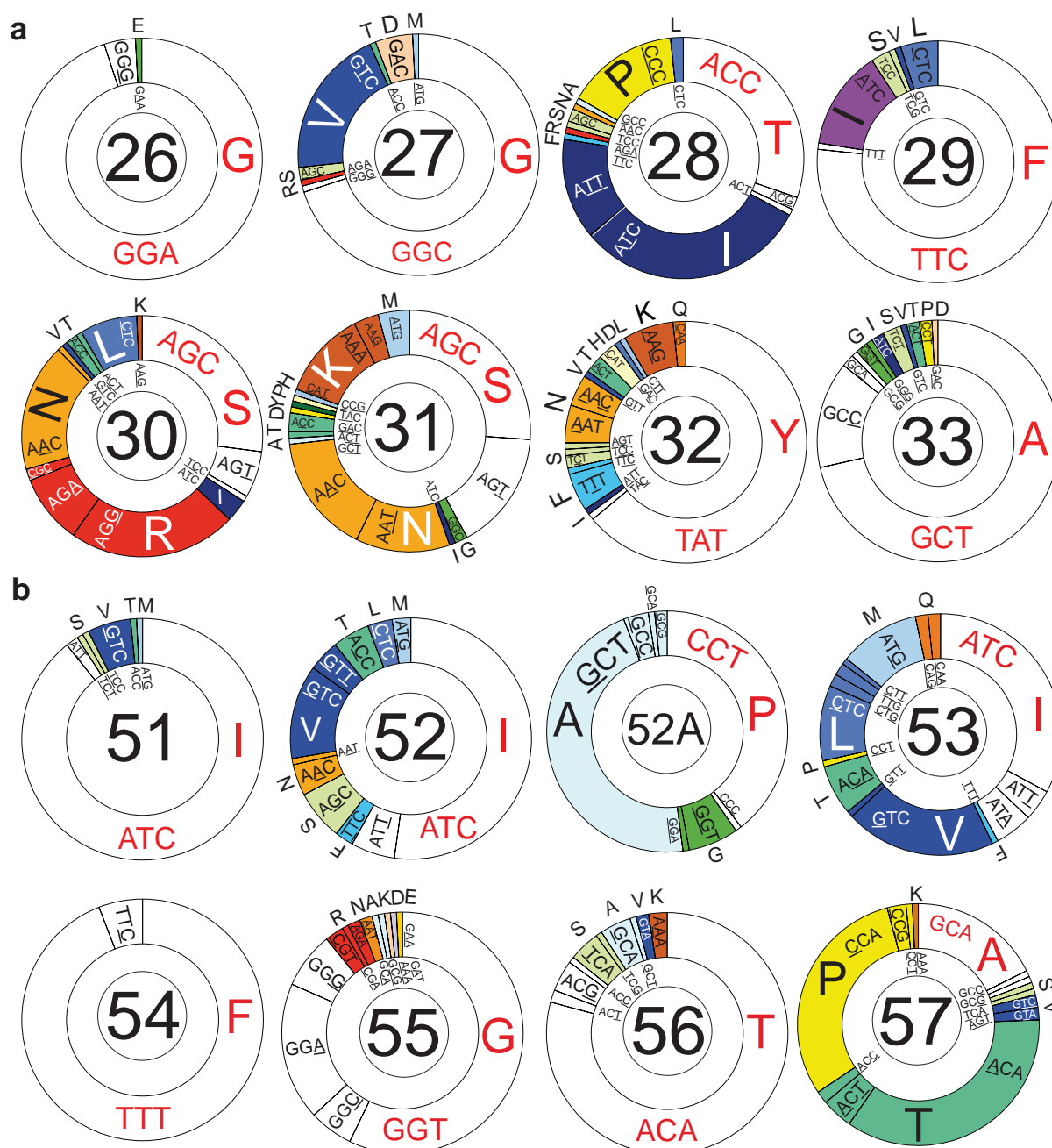


Extended Data Figure 3 | Rapid affinity maturation and accumulation of redundant somatic mutations in the VH1-69 clone 1. **a**, Alignment of VH amino-acid sequences of mutated antibodies with their UCA and branchpoint configurations. The mutations are highlighted in red. Dots indicate identical residues. Residue positions are annotated according to Kabat numbering. **b**, Genealogy tree of clone 1 VH nucleotide sequences. **c**, Binding of

mutated, UCA, branchpoint and variant antibodies to H1 CA09 and H5 VN04 HAs. The mean titre value (\pm s.e.m.) (EC_{50} values of a 1 mg ml^{-1} antibody solution) of three independent experiments is shown. Mutated residues are shown in parentheses. **d**, Neutralization of influenza A viruses. IC_{50} values above $50 \text{ } \mu\text{g ml}^{-1}$ were scored as negative (dotted line). Data are representative of two independent experiments.

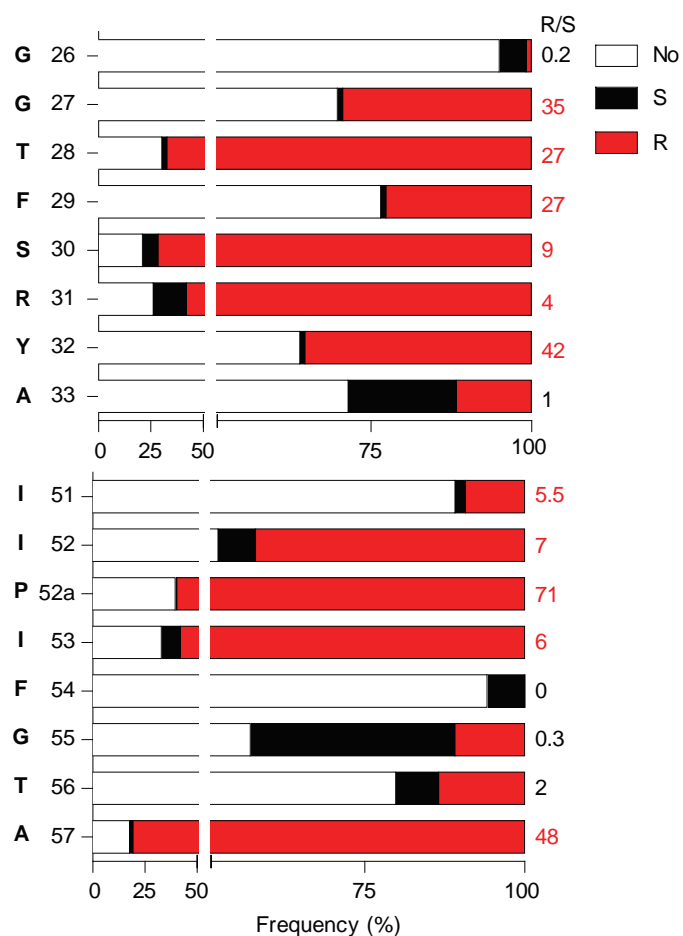


Extended Data Figure 4 | F29 is not required for high affinity binding of VH1-69 antibodies to the HA stem. Binding of mutated and variant antibodies to H1 CA09 HA as measured by ELISA. The mean binding titre value (\pm s.e.m.) (EC_{50} values of a 1 mg ml^{-1} antibody solution) of at least two independent experiments is shown. Mutated residues are shown in parentheses. Data are representative of two independent experiments.

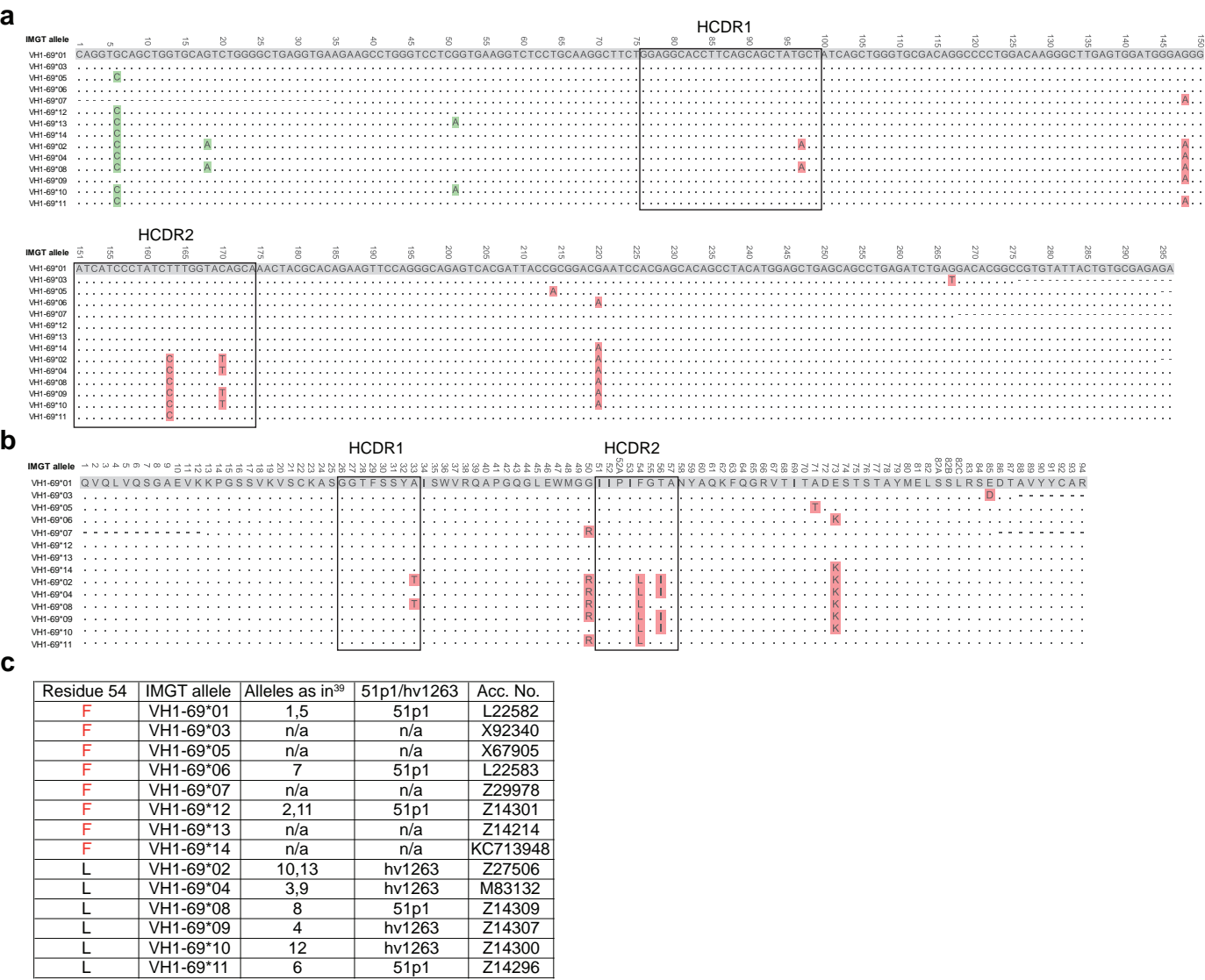


Extended Data Figure 5 | Genetic and mutational requirements in the HCDR1 and HCDR2 of VH1-69 anti-stem antibodies. a, HCDR1; b, HCDR2. Fully annotated pie charts from Fig. 3a, b. Pie charts indicate the

frequency of UCA (white) and mutated residues (IMGT colours), as well as codon usage (underlined nucleotides indicate somatic mutations in that position). UCA codons and amino acids are in red.

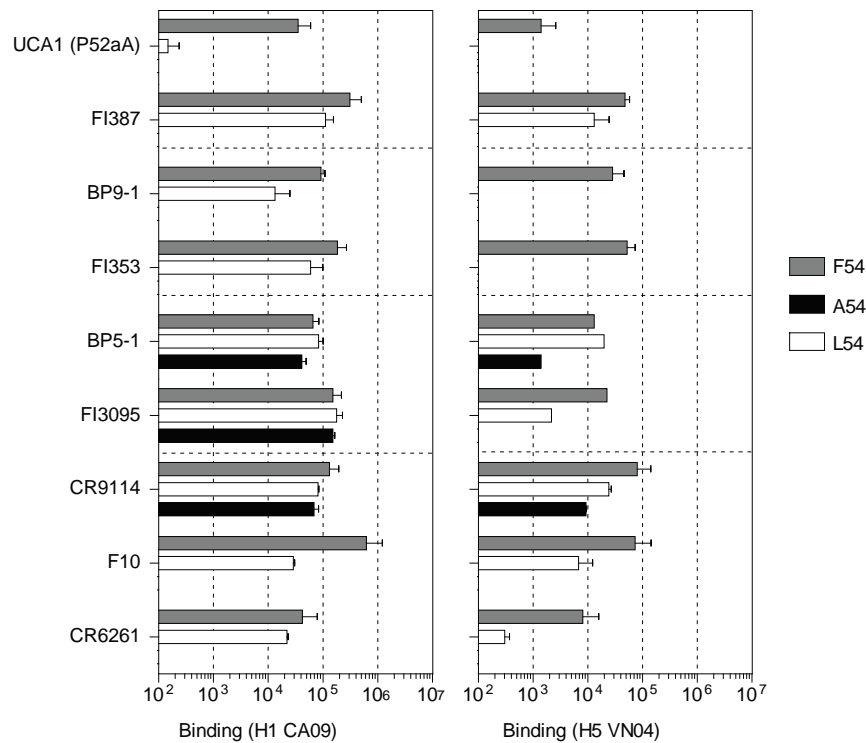


Extended Data Figure 6 | Frequency of silent and replacing mutations at each codon in the HCDR1 and HCDR2 of VH1-69 anti-stem antibodies. The replacement to silent mutation ratio (R/S) values calculated at each codon are shown. Values above 2.9, indicative of positive selection, are highlighted in red.



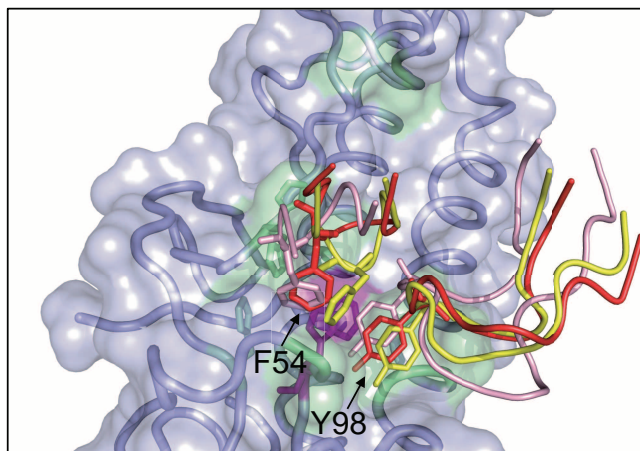
Extended Data Figure 7 | Allelic polymorphism in the *VH1-69* gene.
a, Nucleotide and **b**, amino-acid alignments of the 14 *VH1-69* alleles. Silent mutations are highlighted in green, replacing mutations in red. Dots indicate

identical residues, dashes indicate missing residues in the deposited sequences.
c, Summary table of the different *VH1-69* allele nomenclatures according to their residue at position 54. See also ref. 39.

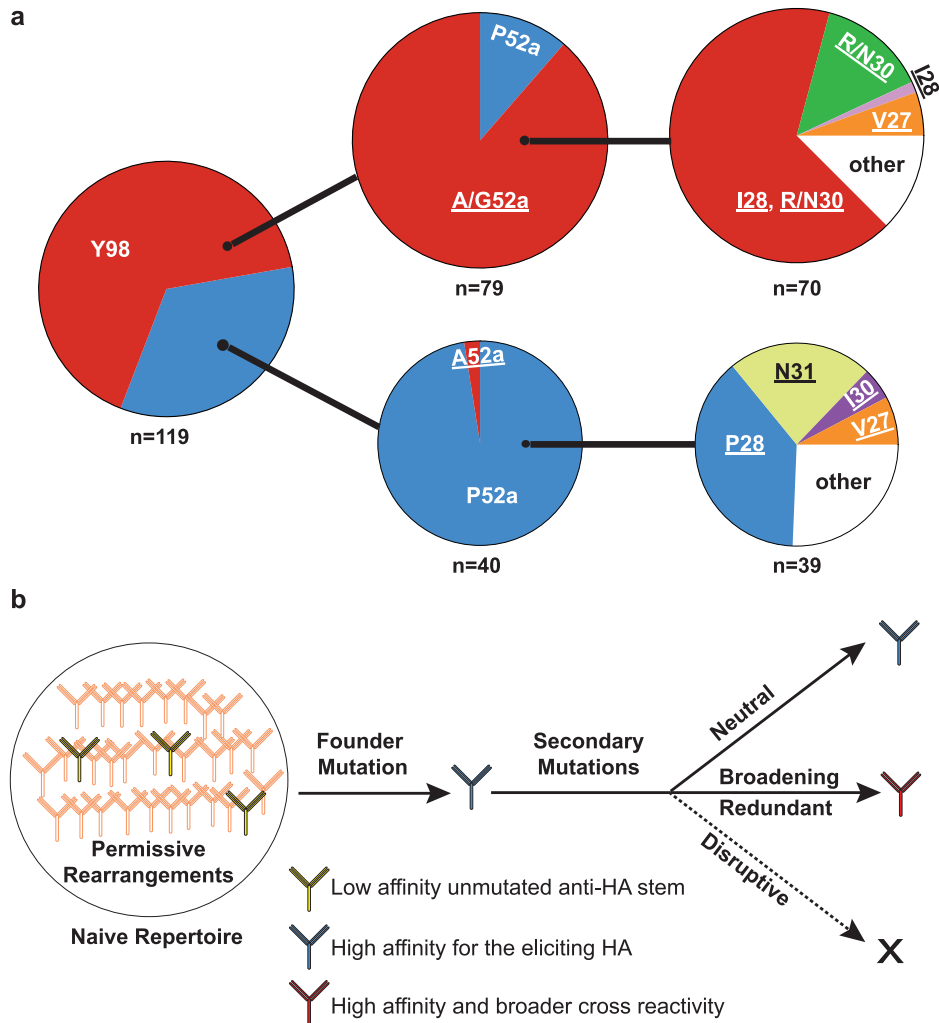


Extended Data Figure 8 | F54 becomes redundant in the context of highly mutated antibodies. Binding of branchpoint and mutated antibodies in their original F54 version (grey), as well as in their A54 (black) or L54 (white)

versions, to H1 CA09 and H5 VN04 HA. The mean ELISA binding titre value (\pm s.e.m.) (EC_{50} values of a 1 mg ml^{-1} antibody solution) of at least two independent experiments is shown.



Extended Data Figure 9 | Structural superimposition of the HCDR2 and HCDR3 of antibodies F10, CR6261 and CR9114. Superimposition of the HCDR2 and HCDR3 of antibodies F10 (pink, PDB 3FKU), CR6261 (red, PDB 3GBM) and CR9114 (yellow, PDB 4FQI) relative to H5 VN04 HA derived by the structural alignment of HA2 atoms on PyMOL, viewed from a different angle compared with Fig. 4c. The antibody F54 and Y98 residues and the conserved HA aromatic residues H18, H38 and W21 are shown as sticks. HCDR2 and HCDR3 amino-acid structural alignments are shown.



Extended Data Figure 10 | Pathways for the development of broadly neutralizing anti-stem VH1-69 antibodies. **a**, The pie charts outline the fraction of antibodies that follow a major pathway characterized by Y98 and a P52aA/G mutation (exemplified by clones 1, 2, 9, 14 and 19) and alternative

pathways found in other clones (exemplified by clone 5). **b**, Schematic representation of the affinity maturation process leading to broadly neutralizing VH1-69 anti-HA stem antibodies.

In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system

Danilo Maddalo¹, Eusebio Manchado¹, Carla P. Concepcion^{1,2}, Ciro Bonetti¹, Joana A. Vidigal¹, Yoon-Chi Han¹, Paul Ogradowski¹, Alessandra Crippa³, Natasha Rekhtman⁴, Elisa de Stanchina⁵, Scott W. Lowe^{1,6} & Andrea Ventura¹

Chromosomal rearrangements have a central role in the pathogenesis of human cancers and often result in the expression of therapeutically actionable gene fusions¹. A recently discovered example is a fusion between the genes echinoderm microtubule-associated protein like 4 (*EML4*) and anaplastic lymphoma kinase (*ALK*), generated by an inversion on the short arm of chromosome 2: inv(2)(p21p23). The *EML4-ALK* oncogene is detected in a subset of human non-small cell lung cancers (NSCLC)² and is clinically relevant because it confers sensitivity to ALK inhibitors³. Despite their importance, modelling such genetic events in mice has proven challenging and requires complex manipulation of the germ line. Here we describe an efficient method to induce specific chromosomal rearrangements *in vivo* using viral-mediated delivery of the CRISPR/Cas9 system to somatic cells of adult animals. We apply it to generate a

mouse model of *Eml4-ALK*-driven lung cancer. The resulting tumours invariably harbour the *Eml4-ALK* inversion, express the *Eml4-ALK* fusion gene, display histopathological and molecular features typical of ALK⁺ human NSCLCs, and respond to treatment with ALK inhibitors. The general strategy described here substantially expands our ability to model human cancers in mice and potentially in other organisms.

Genetically engineered mouse models of human cancers have proven indispensable to dissect the molecular mechanisms underlying tumorigenesis⁴ and provide powerful preclinical platforms for studying drug sensitivity⁵ and resistance^{6–8}. Although many gain- and loss-of-function mutations observed in human cancers can be modelled using current gene-targeting technologies, chromosomal rearrangements leading to oncogenic gene fusions have proven challenging to faithfully recapitulate

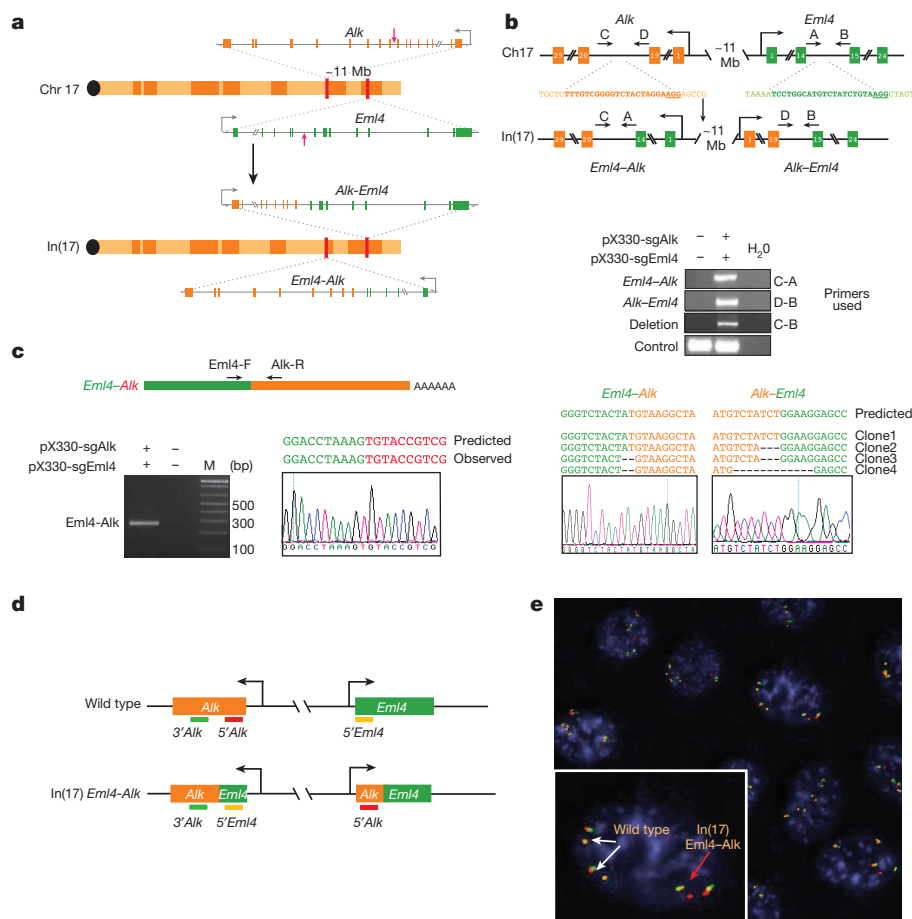


Figure 1 | Induction of *Eml4-ALK* rearrangement in murine cells using the CRISPR-Cas9 system. **a**, Schematic of the In(17) involving the *Eml4* and *Alk* loci. Red arrows indicate the sites recognized by the sgRNAs. **b**, A schematic of the loci before and after the inversion with the location of the primers used (top panel). PCRs were performed on genomic DNA extracted from NIH/3T3 cells transfected with the indicated pX330 constructs (middle panels). The PCR bands were sub-cloned and the sequences of four independent clones and a representative chromatogram are shown in the lower panels. **c**, Schematic of the *Eml4-ALK* fusion transcript by RT-PCR on total RNAs extracted from NIH/3T3 cells transfected with the indicated pX330 constructs (bottom left panel). Sequence of the PCR product showing the correct *Eml4-ALK* junction (bottom right panel). **d**, Schematic of the break-apart interphase FISH strategy. In cells with the *Eml4-ALK* inversion, the red and green probes become separated, and the green and orange probes become juxtaposed. **e**, Break-apart interphase FISH assay on a NIH/3T3 clone selected from cells co-transfected with pX330-*Eml4* and pX330-*Alk*. Both wild type (white arrows) and the In(17) *Eml4-ALK* allele (red arrow) are detected.

¹Memorial Sloan Kettering Cancer Center, Cancer Biology and Genetics Program, 1275 York Avenue, New York, New York 10065, USA. ²Weill Cornell Graduate School of Medical Sciences of Cornell University, 1300 York Avenue, New York, New York 10065, USA. ³Milano-Bicocca University, Department of Medical Oncology, San Gerardo Hospital, 20052, Via G B Pergolesi 33, Monza, Italy. ⁴Memorial Sloan Kettering Cancer Center, Thoracic Pathology and Cytopathology, 1275 York Avenue, New York, New York 10065, USA. ⁵Memorial Sloan Kettering Cancer Center, Molecular Pharmacology Program, 1275 York Avenue, New York, New York 10065, USA. ⁶Howard Hughes Medical Institute, 1275 York Avenue, New York, New York 10065, USA.

in mice. Ectopic expression of fusion oncoproteins from transgenes is widely used to study their oncogenic properties^{9–13}, though with this approach the fusion protein is invariably expressed at non-physiologic levels and neither the role of reduced dosage of the wild-type alleles nor the contribution of the reciprocal product of the translocation can be examined. Strategies that express the fusion transcript from the endogenous locus of the 5' element¹⁴ only partially address these limitations, whereas approaches that engineer *loxP* sites at each breakpoint and produce rearrangements in the presence of Cre recombinase^{15,16} are laborious and have limited applications. Novel genome-editing technologies provide a potentially more flexible strategy to produce precise genomic changes including oncogenic chromosomal rearrangements^{17–20}, but they have not yet been adapted to model such rearrangements *in vivo*.

In the mouse genome, *Eml4* and *Alk* are located on chromosome 17, approximately 11 megabases (Mb) apart, in a region that is syntenic to human chromosome 2(p21–p23) (Fig. 1a). We attempted to model the most common *EML4–ALK* variant in human NSCLCs²¹ by introducing concomitant double-strand DNA breaks at intron 14 of *Eml4* (which corresponds to intron 13 of *EML4*) and at intron 19 of *Alk* (Fig. 1a, b and Extended Data Fig. 1). To induce the DNA breaks we chose the

CRISPR system²² because it only requires co-expression of Cas9 and an appropriately designed single-guide RNA molecule (sgRNA)²³.

We cloned sgRNAs targeting the *Eml4* and *Alk* sites into the Cas9-expressing plasmid pX330 (ref. 24) and co-transfected the resulting constructs in NIH/3T3. PCR analysis demonstrated the induction of the *Eml4–Alk* inversion and of a large deletion of the region between the two cut sites in the transfected cell population (Fig. 1b). The presence of the desired *Eml4–Alk* inversion was confirmed by sequencing the corresponding *Eml4–Alk* fusion transcript (Fig. 1c) and directly visualized by interphase FISH (Fig. 1d, e). Using a similar strategy, we also modelled the *Npm1–Alk* rearrangement, a reciprocal chromosomal translocation commonly observed in anaplastic large cell lymphomas²⁵ (Extended Data Fig. 2). These results confirm that the CRISPR system can be adapted to engineer large deletions, inversions, and chromosomal translocations in eukaryotic cells.

Although appropriate for cell-based experiments, expression of two sgRNAs from separate constructs would be impractical *in vivo*. We therefore engineered plasmids to simultaneously express Cas9 and two distinct sgRNAs from tandem U6 promoters (Extended Data Fig. 3a). Their transfection in NIH/3T3 cells resulted in comparable levels of the two

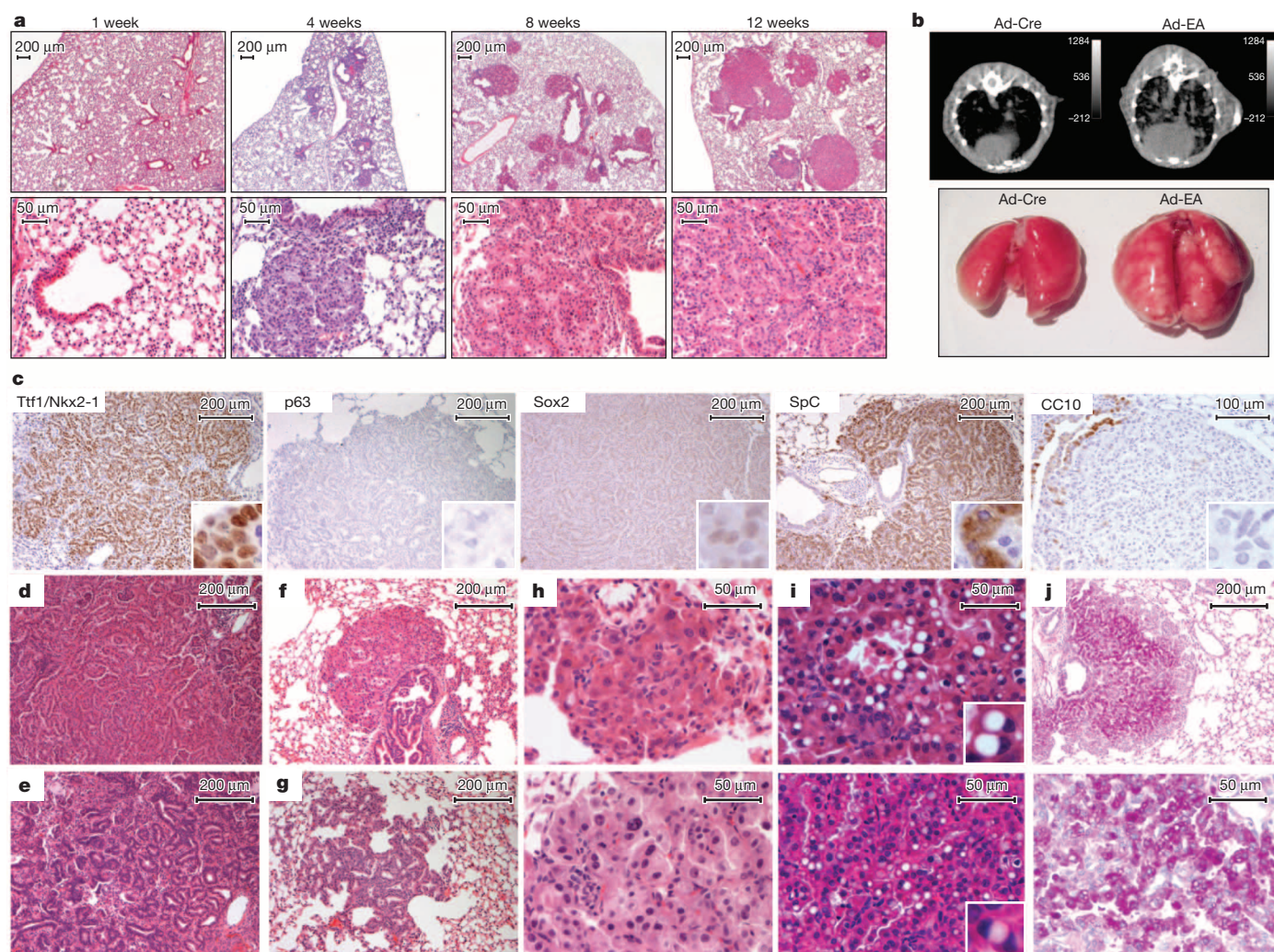


Figure 2 | Intratracheal delivery of Ad-EA leads to lung cancer formation in mice. **a**, Haematoxylin–eosin staining of lungs from mice at the indicated times after intratracheal instillation of Ad-EA. **b**, Representative μ CT scans (top) and macroscopic appearance (bottom) of lungs from mice at 8 weeks post-infection with Ad-Cre or Ad-EA. Numerous neoplastic lesions are evident in the Ad-EA-infected lung. **c**, Representative immunostainings of Ad-EA-induced lung tumours with the indicated antibodies. **d–j**, Tumour architecture

and cytology of Ad-EA-induced tumours. Representative micrographs showing: papillary (**d**) or acinar (**e**) tumours, lesions originating in proximity of intrabronchial hyperplasia (**f**), atypical adenomatous hyperplasia (**g**), mild to moderate nuclear atypia (**h**, top and bottom images), cells with large cytoplasmic vacuole and eccentric nuclei (**i**, top and bottom images), and PAS-positive tumours (**j**, top and bottom images).

sgRNAs, efficient cleavage of the targeted sites, and accumulation of the *Eml4-Alk* inversion (Extended Data Fig. 3b–d).

To deliver Cas9 and sgRNAs targeting the *Alk* and *Eml4* loci to the lungs of adult mice, we next transferred the dual sgRNA/Cas9 cassette into an adenoviral shuttle vector (Extended Data Fig. 4a) and produced recombinant adenoviruses (hereafter referred to as 'Ad-EA'). Adenoviruses are ideal because they efficiently infect the lung epithelium of adult mice²⁶ and do not integrate into the host genome. Infection of mouse embryo fibroblasts (MEFs) with Ad-EA led to the expression of Cas9 and both sgRNAs, and to the rapid generation of the desired *Eml4-Alk* inversion (Extended Data Fig. 4b–d). We estimated that the *Eml4-Alk* inversion was produced in approximately 3–4% of infected MEFs (Extended Data Fig. 4e, f).

To induce the *Eml4-Alk* rearrangement *in vivo* we next infected a cohort of adult CD1 and C57BL/6J (B6) mice by intratracheal instillation of Ad-EA ($n = 52$: 22 B6, 30 CD1) or control adenoviruses expressing either the Cre recombinase (Ad-Cre, $n = 15$: 6 B6, 9 CD1) or Cas9 alone (Ad-Cas9, Fig. 2a–c, $n = 19$: 9 B6, 10 CD1). An annotated list of all infected animals is provided in Extended Data Table 1.

At two days, and at one week post-infection, the lungs appeared histologically normal with no obvious signs of cytotoxicity except for the presence of occasional inflammatory infiltrates (Fig. 2a and data not shown). However, one month after Ad-EA infection, the lungs of mice of both strains presented multiple small lesions that upon histopathological examination appeared to be papillary intrabronchiolar epithelial hyperplasia, atypical adenomatous hyperplasia (AAH) or early well-differentiated adenocarcinomas. By 6–8 weeks post-infection, larger tumours were easily detectable by micro-computed tomography (μ CT) and macroscopically visible at necropsy (Fig. 2b). At 12–14 weeks post-infection, the lungs of Ad-EA-infected mice invariably contained multiple large lesions histologically classified as lung adenocarcinomas.

In Ad-EA-infected animals, multiple bilateral lung tumours were frequently detected by 4–7 weeks post-infection ($n = 23/26$ mice), and invariably after 8 weeks post-infection ($n = 34$). In contrast, Ad-Cre-infected mice remained tumour-free at all time points examined ($n = 14$ mice, range 4–18 weeks), with the exception of two CD1 mice in each of which we observed a single small adenoma. Analogously, even at the latest time point examined (9 weeks post-infection), none of the Ad-Cas9 infected

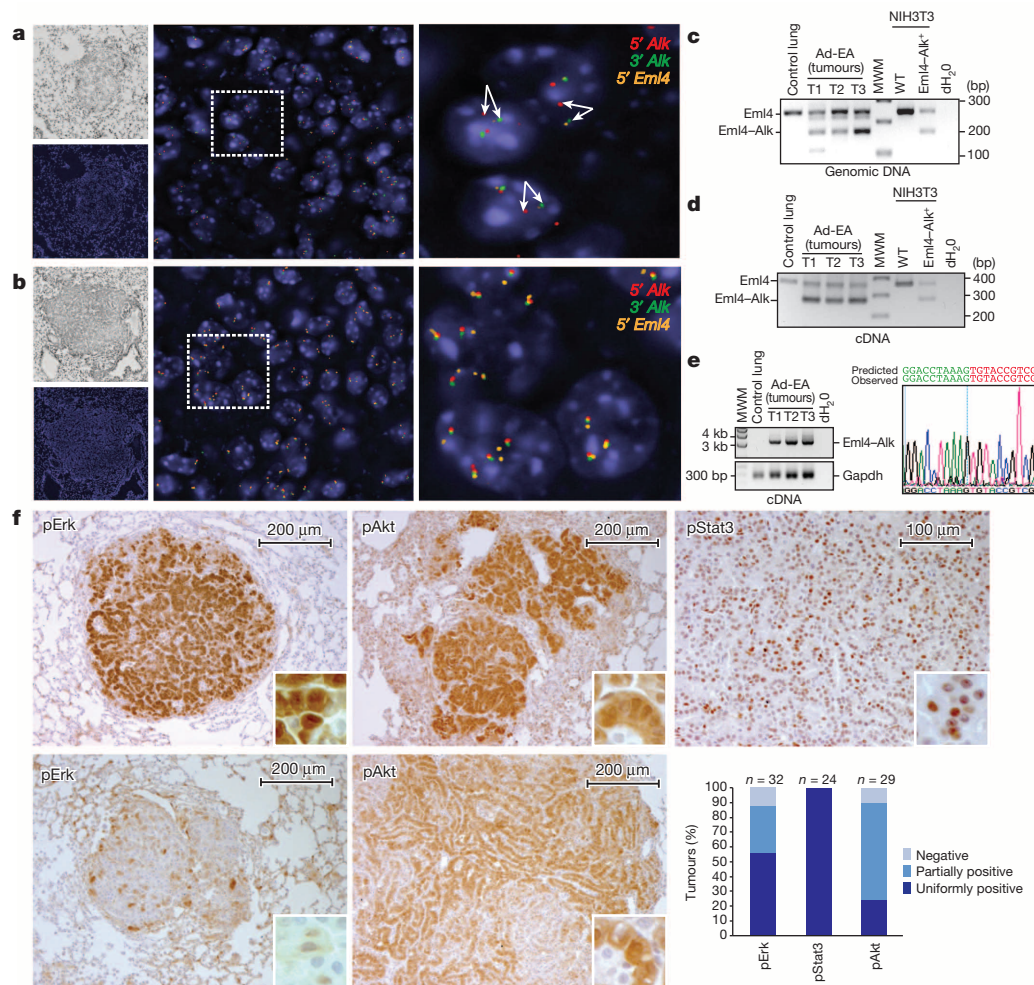


Figure 3 | Lung tumours induced by Ad-EA infection harbour the *Eml4-Alk* inversion. **a, b**, Bright field images and merge fluorescent images at increasing magnification of break-apart interphase FISH showing the presence of the *Eml4-Alk* inversion in a tumour from an Ad-EA-infected mouse (8 weeks post-infection) (**a**) and wild-type configuration of the *Eml4* and *Alk* loci in a control tumour from a conditional K-Ras^{G12D} mouse (**b**). **c**, Detection of the wild-type *Eml4* locus and *Eml4-Alk* inversion in micro-dissected tumours from Ad-EA-infected mice using a three-primer PCR strategy. **d**, RNAs extracted from the same tumours shown in **c** were reverse-transcribed

and amplified using a three-primer strategy to detect the *Eml4* and *Eml4-Alk* transcripts. **e**, RT-PCR detection (left) of the full-length *Eml4-Alk* complementary DNA (~3.2 kilobases (kb)) in the tumours shown in **c**. The full-length PCR products were sequenced on both strands. A chromatogram of the *Eml4-Alk* junction is shown (right). **f**, Representative immunohistochemistry of Ad-EA-induced lung tumours stained with antibodies against the indicated phospho-proteins. A bar-plot of staining intensity for the indicated phospho-proteins is also shown. Tumours from two mice for each group were scored.

mice presented lung tumours ($n = 8$ mice), whereas at same time point all Ad-EA infected mice had developed multiple tumours ($P < 0.0001$, Fisher's exact test). These results indicate that intratracheal delivery of Ad-EA can initiate lung tumorigenesis with high penetrance and low latency, and that this effect cannot be attributed to adenoviral infection or Cas9 expression alone.

All tumours examined were positive for the pneumocyte marker Nkx2-1 (also known as TTF1) and negative for p63 and Sox2, in agreement with the diagnosis of lung adenocarcinoma (Fig. 2c). The tumours were also strongly positive for the alveolar type II marker surfactant protein C (SpC), whereas the Clara cell marker CCSP (also known as CC10) was undetectable. The adenocarcinomas had a papillary or, less frequently, acinar architecture (Fig. 2d, e). Most of these tumours were in close proximity to bronchi and bronchioles showing papillary epithelial hyperplasia (Fig. 2a, f), and areas of AAH were frequently observed, especially at earlier time points (Fig. 2g). The majority of tumour cells appeared low-grade, with occasional instances of intermediate nuclear atypia with enlarged nuclei and prominent nucleoli (Fig. 2h, top and bottom images). Approximately 20% of tumours contained cells with a large cytoplasmic vacuole and a peripherally located nucleus (Fig. 2i, top and bottom images). These cells are reminiscent of signet ring cells, which are commonly observed in human ALK⁺ NSCLC²⁷. Approximately 30% of adenocarcinomas displayed areas of intense positivity at the periodic acid-Schiff (PAS) staining (Fig. 2j, top and bottom images).

Interphase FISH analysis demonstrated the presence of a mono- or bi-allelic *Eml4-Alk* inversion in every Ad-EA-induced tumour examined ($n = 4$ animals) (Fig. 3a), but not in control K-Ras^{G12D}-driven tumours²⁸ (Fig. 3b). We further confirmed the presence of the *Eml4-Alk* rearrangement and expression of the full-length *Eml4-Alk* transcript in microdissected tumours by performing genomic PCR and reverse transcription PCR followed by sequencing (Fig. 3c–e).

Activation of the human *ALK* oncogene via deregulation, translocation, or amplification has been shown to lead to constitutive phosphorylation of ERK, STAT3, and AKT²⁹. At 12–14 weeks post-infection, all lung tumours derived from Ad-EA-injected mice showed phosphorylation and nuclear localization of Stat3. Phosphorylation of Akt and Erk1/2 were also frequently, but not invariably, observed (Fig. 3f).

Finally, we examined the sensitivity of Ad-EA-induced lung tumours to crizotinib, a dual ALK/MET inhibitor used in the clinic to treat patients affected by ALK⁺ NSCLCs³. Ten Ad-EA-infected CD1 mice were monitored by μ CT scans starting at 9 weeks post-infection until the appearance of multiple large lung tumours, at which point the animals were randomly assigned to receive a daily dose of crizotinib ($n = 7$) or vehicle ($n = 3$) (Fig. 4a). After two weeks of treatment the animals in the crizotinib group displayed complete (6/7) or partial (1/7) tumour regression, as indicated by μ CT scans and confirmed at necropsy, whereas all control animals showed signs of disease progression (Fig. 4b, c, Extended Data Fig. 5, Extended Data Table 2 and Supplementary Videos 1–10). Histological analysis showed that in the crizotinib group the tumours had undergone marked atrophy or were replaced by areas of intense inflammatory necrosis (Fig. 4d, e).

Collectively, these results demonstrate that the CRISPR technology can be adapted to engineer oncogenic chromosomal rearrangements in mice. The new mouse model of *Eml4-Alk*-driven lung cancer we have generated to validate this approach faithfully recapitulates the molecular and biological properties of human ALK⁺ NSCLCS, including a marked sensitivity to the ALK-inhibitor crizotinib. This model provides unique opportunities to dissect the molecular mechanisms through which *Eml4-Alk* drives tumour formation, to test the efficacy of targeted therapies, and to investigate the mechanisms of drug resistance *in vivo*.

The CRISPR-based strategy described here offers several advantages over germline engineering via transgenesis or homologous recombination.

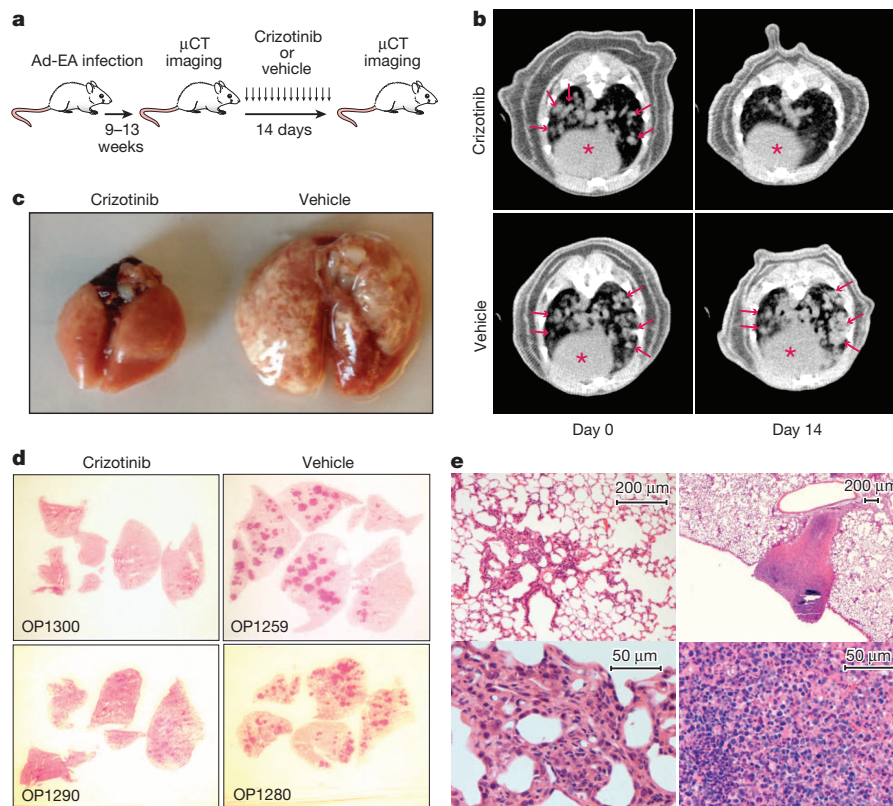


Figure 4 | Ad-EA-induced lung tumours respond to crizotinib treatment. **a**, Schematic of the experiment. **b**, Representative μ CT of the lungs of mice treated with crizotinib or vehicle at day 0 and after 2 weeks of treatment. Lung tumours are indicated by arrows. Red asterisks mark the hearts. **c**, Macroscopic appearance of the lungs after 2 weeks of treatment. **d**, Low

magnification of lung sections from two crizotinib-treated and two vehicle-treated mice (haematoxylin and eosin). **e**, Higher magnification of representative haematoxylin and eosin stained lung sections from crizotinib-treated mice showing residual atrophic foci of tumour cells (left) or necrotic-inflammatory debris (right).

By inducing the rearrangement in only a subset of somatic cells, the resulting lesions more closely recapitulate the stochastic nature of tumour formation in humans. In addition, by modifying the endogenous loci, expression of the resulting fusion genes is subjected to physiologic transcriptional and post-transcriptional regulation, accurately modelling the reduced dosage of the wild-type alleles and the expression of the reciprocal product of the translocation/inversion. Finally, because our method requires only the generation of an appropriate viral vector and no germline manipulations, it can be readily adapted to model chromosomal rearrangements in other species, including non-human primates, and as such will facilitate the study of species-specific differences in tumour progression and therapy response *in vivo*.

Despite these key advantages, some caveats of the CRISPR technology must also be considered. The efficiency with which the rearrangements are induced is relatively low and is likely to be affected by the distance between the cut sites and their accessibility to Cas9. Although a low efficiency may be desirable when inducing oncogenic rearrangements, it is a concern if the goal is to generate chromosomal rearrangements in the majority of cells. Furthermore, every possible allele combination of the two target loci (indels, inversions, deletions) will be induced by the dual sgRNA/Cas9 system³⁰, potentially complicating the interpretation of these studies.

In summary, the general strategy we have developed substantially expands our ability to model cancers driven by chromosomal rearrangements and will facilitate the development of pre-clinical models to study the mechanisms of drug resistance and test novel therapies.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 28 July; accepted 30 September 2014.

Published online 22 October 2014.

1. Taki, T. & Taniwaki, M. Chromosomal translocations in cancer and their relevance for therapy. *Curr. Opin. Oncol.* **18**, 62–68 (2006).
2. Soda, M. *et al.* Identification of the transforming *EML4-ALK* fusion gene in non-small-cell lung cancer. *Nature* **448**, 561–566 (2007).
3. Kwak, E. L. *et al.* Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N. Engl. J. Med.* **363**, 1693–1703 (2010).
4. Tuveson, D. A. & Jacks, T. Technologically advanced cancer modeling in mice. *Curr. Opin. Genet. Dev.* **12**, 105–110 (2002).
5. Sharpless, N. E. & Depinho, R. A. The mighty mouse: genetically engineered mouse models in cancer drug development. *Nature Rev. Drug Discov.* **5**, 741–754 (2006).
6. Pirazzoli, V. *et al.* Acquired resistance of EGFR-mutant lung adenocarcinomas to afatinib plus cetuximab is associated with activation of mTORC1. *Cell Rep.* **7**, 999–1008 (2014).
7. Bergers, G. & Hanahan, D. Modes of resistance to anti-angiogenic therapy. *Nature Rev. Cancer* **8**, 592–603 (2008).
8. Rottenberg, S. *et al.* Selective induction of chemotherapy resistance of mammary tumors in a conditional mouse model for hereditary breast cancer. *Proc. Natl Acad. Sci. USA* **104**, 12117–12122 (2007).
9. Heisterkamp, N. *et al.* Acute leukaemia in *bcr/abl* transgenic mice. *Nature* **344**, 251–253 (1990).
10. Zuber, J. *et al.* Mouse models of human AML accurately predict chemotherapy response. *Genes Dev.* **23**, 877–889 (2009).
11. Lange, K. *et al.* Overexpression of NPM-ALK induces different types of malignant lymphomas in IL-9 transgenic mice. *Oncogene* **22**, 517–527 (2003).
12. Chiarle, R. *et al.* NPM-ALK transgenic mice spontaneously develop T-cell lymphomas and plasma cell tumors. *Blood* **101**, 1919–1927 (2003).
13. Soda, M. *et al.* A mouse model for *EML4-ALK*-positive lung cancer. *Proc. Natl Acad. Sci. USA* **105**, 19893–19897 (2008).
14. Corral, J. *et al.* An *Mll-AF9* fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes. *Cell* **85**, 853–861 (1996).
15. Smith, A. J. *et al.* A site-directed chromosomal translocation induced in embryonic stem cells by *Cre-loxP* recombination. *Nature Genet.* **9**, 376–385 (1995).
16. Collins, E. C., Pannell, R., Simpson, E. M., Forster, A. & Rabbitts, T. H. Inter-chromosomal recombination of *Mll* and *Af9* genes mediated by *cre-loxP* in mouse development. *EMBO Rep.* **1**, 127–132 (2000).
17. Piganeau, M. *et al.* Cancer translocations in human cells induced by zinc finger and TALE nucleases. *Genome Res.* **23**, 1182–1193 (2013).
18. Torres, R. *et al.* Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR-Cas9 system. *Nature Commun.* **5**, 3964 (2014).
19. Brunet, E. *et al.* Chromosomal translocations induced at specified loci in human stem cells. *Proc. Natl Acad. Sci. USA* **106**, 10620–10625 (2009).
20. Choi, P. S. & Meyerson, M. Targeted genomic rearrangements using CRISPR/Cas technology. *Nature Commun.* **5**, 3728 (2014).
21. Choi, Y. L. *et al.* Identification of novel isoforms of the *EML4-ALK* transforming gene in non-small cell lung cancer. *Cancer Res.* **68**, 4971–4976 (2008).
22. Horvath, P. & Barrangou, R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167–170 (2010).
23. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
24. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
25. Morris, S. W. *et al.* Fusion of a kinase gene, *ALK*, to a nucleolar protein gene, *NPM*, in non-Hodgkin's lymphoma. *Science* **263**, 1281–1284 (1994).
26. DuPage, M., Dooley, A. L. & Jacks, T. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nature Protocols* **4**, 1064–1072 (2009).
27. Nishino, M. *et al.* Histologic and cytomorphologic features of *ALK*-rearranged lung adenocarcinomas. *Modern Pathol.* **25**, 1462–1472 (2012).
28. Jackson, E. L. *et al.* Analysis of lung tumor initiation and progression using conditional expression of oncogenic *K-ras*. *Genes Dev.* **15**, 3243–3248 (2001).
29. Chiarle, R., Voena, C., Ambrogio, C., Piva, R. & Inghirami, G. The anaplastic lymphoma kinase in the pathogenesis of cancer. *Nature Rev. Cancer* **8**, 11–23 (2008).
30. Canver, M. C. *et al.* Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *J. Biol. Chem.* **289**, 21312–21324 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank M. Fazio, M. Ladanyi, G. Riely, S. Armstrong, and the members of the Ventura, Lowe and Jacks laboratories for discussion and comments. We also thank J. Hollenstein for editing the manuscript, T. Jacks for providing tumour samples from K-Ras^{G12D} mice, and the Cytogenetic Core Facility of MSKCC for tissue processing and histology. This work was supported by grants from the Geoffrey Beene Cancer Research Foundation (A.V.), NCI (Cancer Center Support Grant P30 CA008748, E.d.S.), HHMI (S.W.L.), NCI Project Grant (S.W.L.); and by fellowships from the American Italian Cancer Foundation (D.M.), the Foundation Blanche Boncompagni Ludovisi, née Bildt (D.M.), and the Jane Coffin Childs Foundation (E.M.). C.P.C. was supported by an NCI training grant.

Author Contributions D.M. and A.V. conceived the project, designed and analysed the experiments, and wrote the manuscript. S.W.L. contributed to the interpretation of the results and the writing of the manuscript. D.M. generated and tested the constructs, performed the cell-based experiments, and characterized the *Eml4-Alk* tumours. E.M., D.M., C.B., Y.-C.H. and P.O. performed the *in vivo* experiments. E.d.S. supervised the crizotinib treatment experiments and analysed the results. J.A.V., D.M., C.P.C. and A.V. microdissected and analysed lung tumours to detect the *Eml4-Alk* inversion. C.B., D.M. and A.C. performed the immunostainings. N.R. reviewed the histopathology.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.V. (venturaa@mskcc.org).

METHODS

Plasmids and adenoviral vectors. The pX330 vector expressing Cas9 (Addgene plasmid 42230) was digested with BbsI and ligated to annealed and phosphorylated sgRNA oligonucleotides targeting *Eml4*, *Alk* and *Npm1*. For cloning of tandem U6-sgRNA-Cas9 constructs, the second U6-sgRNA cassette was amplified using primers containing the XbaI and KpnI sites and cloned into the pX330 construct containing the appropriate sgRNA. For Adeno-*Eml4*-*Alk* cloning, pX330-*Alk*-*Eml4* vector was modified by adding an XhoI site upstream the first U6 promoter. An EcoRI/XhoI fragment containing the double U6-sgRNA cassette and the Flag-tagged Cas9 was then ligated into the EcoRI/XhoI-digested pacAd5 shuttle vector. NIH/3T3 cells were transfected in 6-well plates with 3 µg of total plasmid DNA per well using lipofectamine 2000 (Invitrogen) following the manufacturer's instructions. To enrich for transfected cells, transfections included 1 µg of a plasmid expressing the Puro-resistance gene (pSico) and cells were incubated with 2 µg ml⁻¹ puromycin for 2 days. Recombinant adenoviruses were generated by Viraplex (Ad-EA and Ad-Cas9) or purchased from the University of Iowa (Ad-Cre). MEFs infections were performed by adding adenovirus (3 × 10⁶ plaque-forming units (p.f.u.)) to each well of a 6-well plate.

PCR and RT-PCR analysis. For PCR analysis of genomic DNA, cells were collected in lysis buffer (100 mM Tris-HCl pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl supplemented with fresh proteinase K at final concentration of 100 ng ml⁻¹). Genomic DNA was extracted with phenol-chloroform-isoamyl alcohol and precipitated in ethanol. The DNA pellet was dried and resuspended in double-distilled water. For RT-PCR, total RNAs were extracted with TRIzol (Life Technologies) following manufacturer's instructions. cDNAs were prepared using the Superscript III kit, following the manufacturer's instructions. The primers and the primer pairs used in the various PCR reactions are provided in Extended Data Tables 3 and 4.

Quantification of inversion efficiency in MEFs. We first isolated an NIH/3T3 subclone carrying a mono-allelic *Eml4*-*Alk* inversion validated by interphase FISH. Genomic DNA extracted from this clone was mixed with increasing amounts of genomic DNA from parental NIH/3T3 cells to generate a series of standards containing known percentage of *Eml4*-*Alk* alleles. The standards and the test samples were then subjected to quantitative PCR (Applied Biosystem) using primers amplifying the *Eml4*-*Alk* junction (*Eml4*-for and *Alk*-rev, see Extended Data Table 3) or a control gene (*miR*-17-92-gDNA-for and *miR*-17-92-gDNA-rev) and the fraction of *Eml4*-*Alk* alleles in the test was calculated by plotting the $\Delta\Delta C_t$ values on the standard curve. qPCR analysis was performed using SYBR Green (Life Technology).

Cell lines. MEFs were generated from E14.5 wild-type embryos following standard procedures. NIH/3T3 were purchased from ATCC.

Mouse husbandry and adenoviral infection. Mice were purchased from The Jackson Laboratory (C57BL/6J) or from Charles River (CD1) and housed in the SPF MSKCC animal facility, where the health status of the colony is constantly monitored by the veterinary staff and by a sentinel program. For adenoviral infection, 6–10-week-old mice were anaesthetized by intraperitoneal injection of ketamine (80 mg per kg) and xylazine (10 mg per kg) and treated by intratracheal instillation of 1.5 × 10⁸ p.f.u. adenovirus per mouse, as previously described²⁶. Investigators were not blinded with respect to which adenovirus was injected. All studies and procedures were approved by the Memorial Sloan-Kettering Cancer Center Institutional Animal Care and Use Committee.

Interphase fluorescent *in situ* hybridization. Interphase FISH experiments were performed and interpreted by the MSKCC cytogenetic core using a 3-colour probe

mix designed to detect and discriminate between *Alk*-*Eml4* fusion and other rearrangements of *Alk*. The probe mix comprised mouse BAC clones mapping to: 3' *Alk* (17qE1.3, RP23-306H20, RP23-397M18 labelled with green dUTP), 5' *Alk* (17qE1.3, RP23-12H17, RP23-403F20 labelled with red dUTP), and 5' *Eml4* (17qE4, RP23-193B15 labelled with orange dUTP). Probe labelling, hybridization, washing, and fluorescence detection were done according to standard procedures. Cell line collection and metaphase spreads were prepared according to standard cytogenetics procedures. For NIH/3T3, FISH signals were enumerated in a minimum of 20 metaphases to determine locus specificity, and 100 interphase cells to determine *Alk*-*Eml4* fusion status. Each paraffin section was first scanned under ×100 objective to assess signal pattern and select representative regions for analysis. At least three images per representative region were captured (each image was a compressed stack of 12 z-sections at 0.5 micron intervals). Signal counts were performed on the captured images and a minimum of 50 interphase nuclei was analysed to determine the *Alk*-*Eml4* fusion status. Based on the observed distance between the green (3' *Alk*), red (5' *Alk*), and orange (5' *Eml4*) signal in the negative controls (parental cell line and Ad-Cre-infected cells), interphase cells were classified as normal, *Eml4*-*Alk* positive, or other.

Surveyor assay. The genomic region flanking the CRISPR/Cas9 target site was first amplified by PCR. After a cycle of melting and re-annealing to allow heteroduplex formation, the amplicon was digested with the surveyor nuclease (Transgenomic) for 1 h at 42 °C according to the manufacturer's directions and the digestion products were separated on a 2% agarose gel.

Northern blot analysis. 10 µg of RNA previously extracted with TRIzol (Life Technologies) were run on a 15% denaturing polyacrylamide gel and blotted on a nitrocellulose membrane for 1 h at 100 V at room temperature. The membranes were then hybridized to radiolabelled oligonucleotides complementary to the *Alk* (5'-TACAGATAGACATGCCAGGAC), *Eml4* (5'-TCCTAGTAGACCCGACAAAC) sgRNAs, or mU6 (5'-GCAGGGGCCATGCTAATCTTCTGTATCG) dissolved in ExpressHyb (Clontech) at 42 °C overnight. Washes were performed at room temperature in 2X SSC and 0.2 SSC.

Lung processing and antibodies for immunohistochemistry. Lungs were inflated by intratracheal injection of 4% paraformaldehyde (PFA), incubated for 18–24 h in 4% PFA, and then transferred to 70% ethanol for at least 24 h before further processing. The following antibodies were used: phospho-Stat3 (Tyr705, Cell Signaling Technology #9135, 0.1 µg ml⁻¹); phospho-Erk1/2 (Thr202/Tyr204, Cell Signaling Technology #4370 1 µg ml⁻¹); phospho-Akt (Ser473, Cell Signaling Technology #4060 1 µg ml⁻¹); Nkx2-1 (Epitomics, EP1584Y 1:1,200); Flag (Sigma, M2 1:1,000); P63 (Santa Cruz (H-137) sc8343, 1:1,000); Sox2 (Cell Signaling Technology, C70B1 #3728, 1:1,000); CC10/CCSP (Millipore, 07-623, 1:2,000); SpC (Millipore, AB3786, 1:1,000).

µCT imaging. µCT Scans were performed on the Mediso Nano SPECT/CT System covering only the lung fields of each mouse. Each scan averaged approximately 5 min using 240 projections with an exposure time of 1,000 ms set at a pitch of 1 degree. The tube energy of the X-ray was 55 kVp and 145 µA. The in-plane voxel sizes chosen were small and thin creating a voxel size of 73 × 73 × 73 µm. The final reconstructed image consisted of 368 × 368 × 1,897 voxels. Scans were analysed with the Osirix software.

Crizotinib treatment. Mice were randomized to receive either control vehicle (water) or crizotinib at 100 mg per kg per os daily for at least 14 consecutive days. Mice were monitored daily for weight loss and clinical signs. Investigators were not blind with respect to treatment.

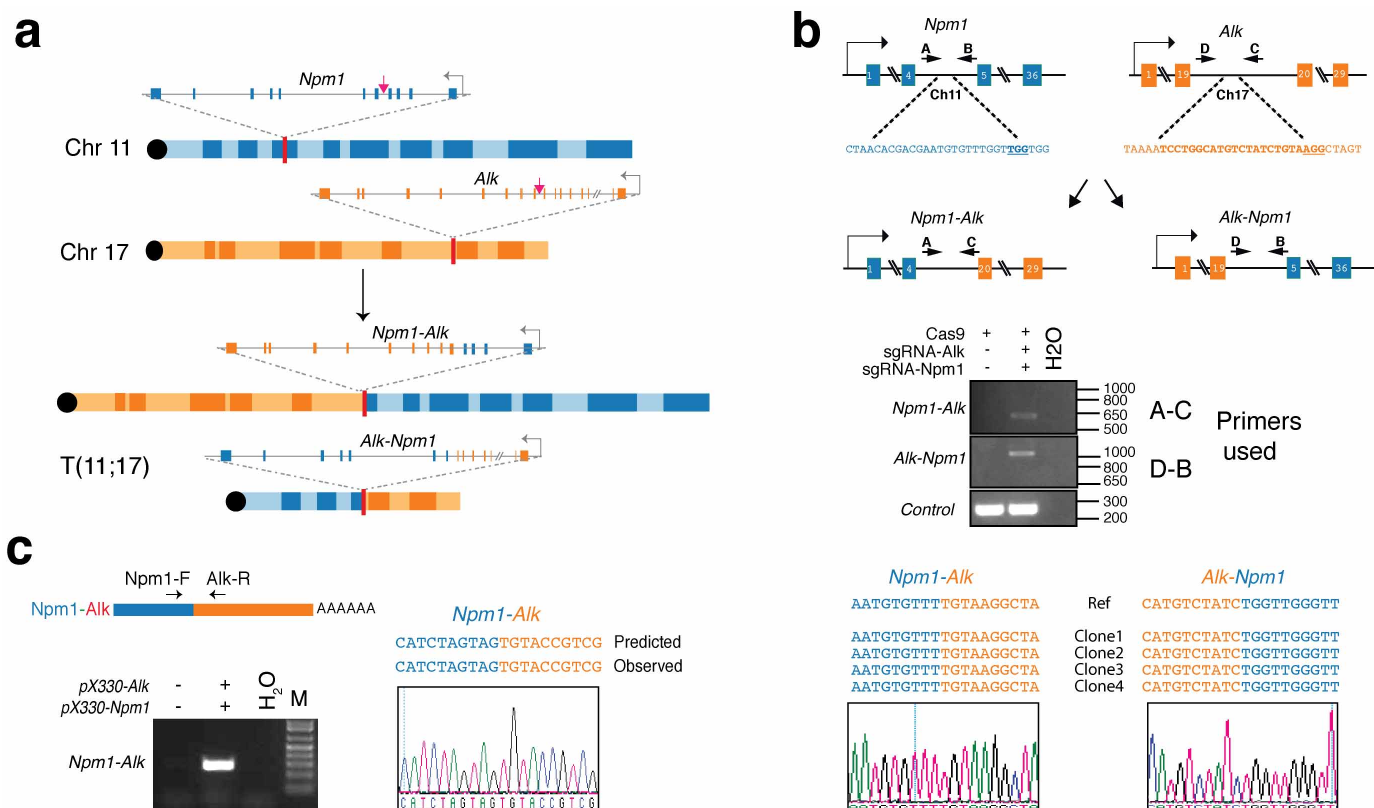
a

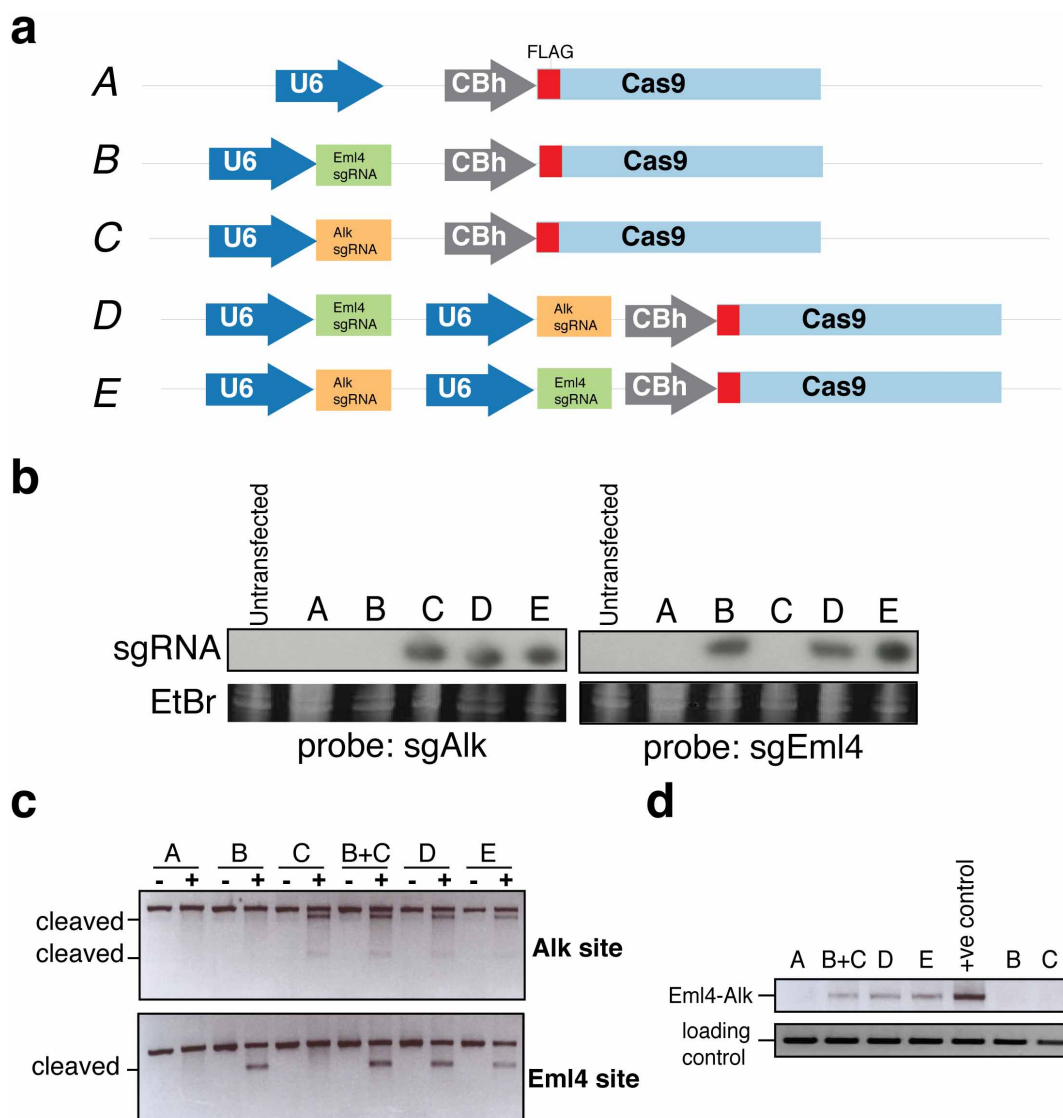
ENSMUST00000096766_Exon_14 (Eml4)	AAATATGAAAAACCAAAATTCGTTCACTGTTTGGCATTCTTGGGGAATGG 50
ENST00000318522_Exon_13 (EML4)	AAATATGAAAAGCCAAAATTTGTGCAGTGTTCAGTATTCTTGGGGAATGG 50
	***** ** *****
ENSMUST00000096766_Exon_14 (Eml4)	AGATGTTCTCACTGGAGACTCGGGTGGAGTCATGCTGATCTGGAGCAAAA 100
ENST00000318522_Exon_13 (EML4)	AGATGTTCTTACTGGAGACTCAGGTGGAGTCATGCTTATATGGAGCAAAA 100
	***** ** *****
ENSMUST00000096766_Exon_14 (Eml4)	CGATGGTAGAGCCCCCGCCGGGAAAGGACCTAAAG 136
ENST00000318522_Exon_13 (EML4)	CTACTGTAGAGCCACACCTGGGAAAGGACCTAAAG 136
	* * ***** * ** *****

b

	EML4	ALK	
EML4-ALK (variant 1)	...IWSKTTVEPTPGKGP	KVYRRKHQELQAMQMELQ...	(HUMAN)
Eml4-Alk (this study)	...IWSKTMVEPPP	GP	KVYRRKHQELQAMQIQLO...
	Em14	Alk	

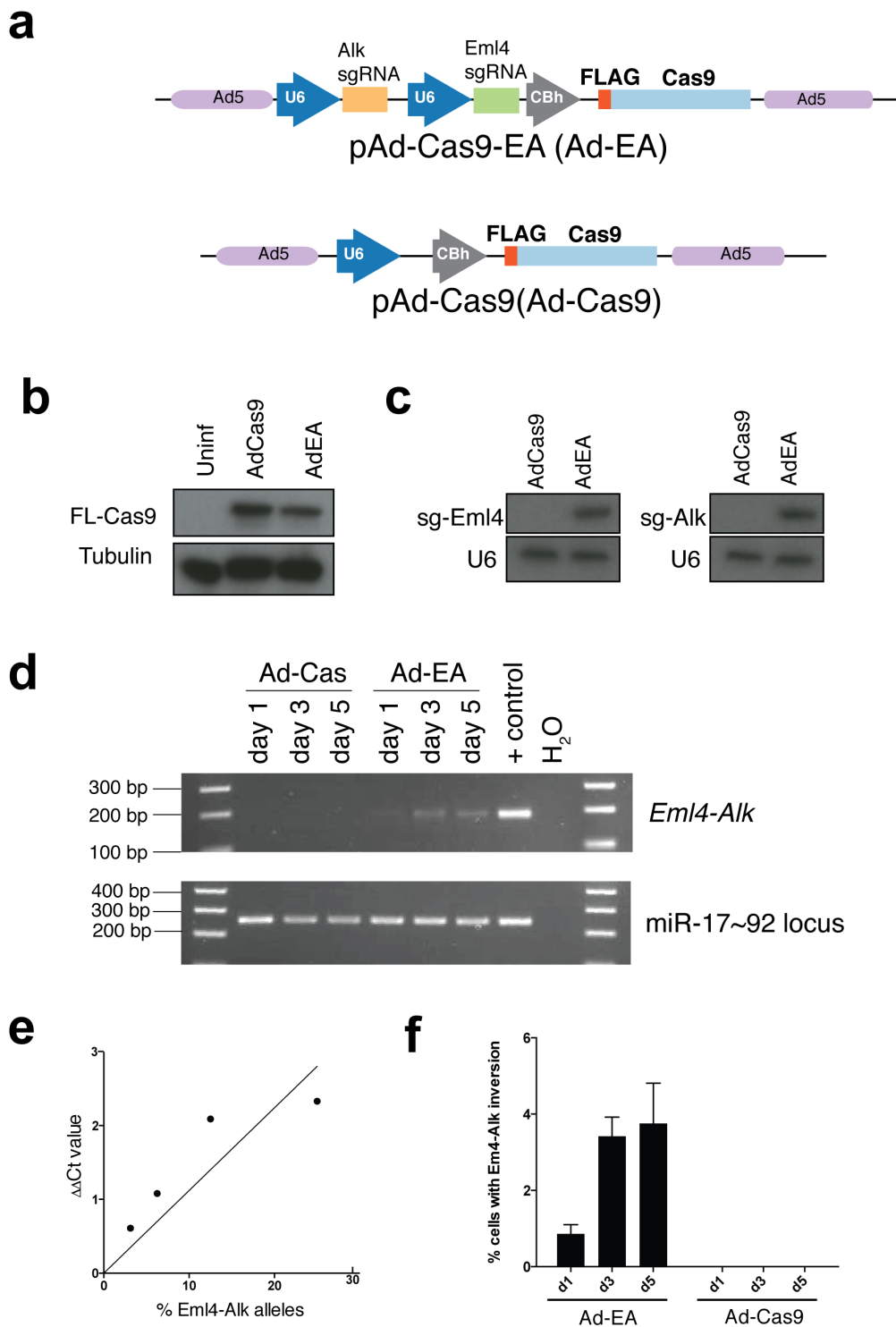
Extended Data Figure 1 | Human and murine Eml4–Alk. **a**, Alignment of human EML4 exon 13 and mouse Eml4 exon 14. **b**, Alignment of the junction between the human EML4–ALK (variant 1) and the predicted murine Eml4–Alk proteins.





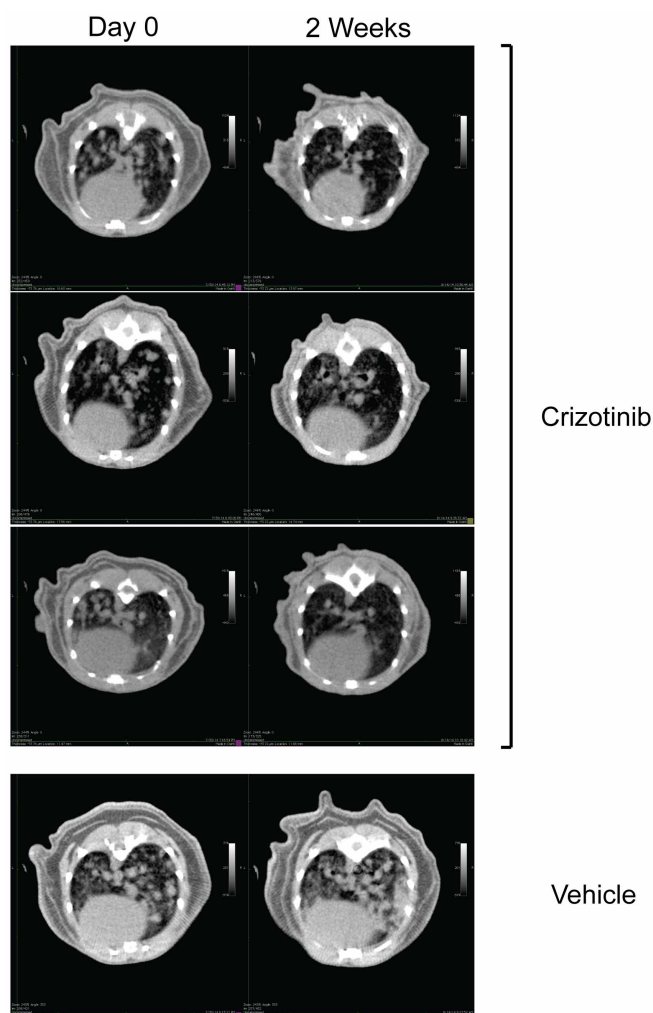
Extended Data Figure 3 | Comparison of dual and single sgRNA-expressing plasmids. **a**, Schematic of pX330 (A) and its derivatives (B–E) used in these experiments. NIH/3T3 were transfected with these constructs and lysed to extract total RNA and genomic DNA. **b**, RNAs were analysed by northern

blotting with probes against the *Alk* (left) or *Eml4* (right) sgRNAs. **c**, **d**, The DNA samples were subjected to surveyor assays (**c**), or amplified by PCR to detect the *Eml4*–*Alk* inversion (**d**).



Extended Data Figure 4 | Induction of the *Eml4-Alk* inversion in primary MEFs using an adenoviral vector expressing Flag-Cas and tandem sgRNAs. **a**, Schematic of the adenoviral vectors. **b**, Immunoblot using an anti-Flag antibody on lysates from MEFs infected with the indicated adenoviruses. **c**, Small-RNA northern blots using probes against sgEml4 and sgAlk on total RNAs from cells infected with Ad-Cas9 or Ad-EA. **d**, PCR-mediated detection of the *Eml4-Alk* inversion in MEFs infected with Ad-Cas9 or Ad-EA for the

indicated number of days. **e**, Standard curve generated performing quantitative PCR analysis on genomic DNA containing a known fraction of *Eml4-Alk* alleles. Average of two independent experiments. **f**, Quantification of the fraction of MEFs harbouring the *Eml4-Alk* inversion at the indicated time points after infection with Ad-EA or Ad-Cas9. Values are mean of three independent infections \pm s.d.



Extended Data Figure 5 | Radiologic response of Ad-EA-induced tumours to crizotinib treatment. μ CT images from crizotinib- or vehicle-treated mice at day 0 and after 2 weeks of treatment.

Extended Data Table 1 | Mouse cohorts

				Time since infection (weeks)																			
Mouse	Sex	Virus	Strain	Notes	2days	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
OP1925	M	Ad-CRE	B6		NO †																		
OP1110	F	Ad-CRE	B6						NO †														
notag-1	F	Ad-CRE	CD1	A single small adenoma was observed on tissue sections					YES †														
OP1285	F	Ad-CRE	CD1																				
OP1254	F	Ad-CRE	CD1											NO †									
OP1106	F	Ad-CRE	B6								NO #				NO #		NO †						
OP1276	F	Ad-CRE	CD1							NO #												NO #	
OP1116	F	Ad-CRE	B6								NO #				NO #							NO #	
OP1284	F	Ad-CRE	CD1								NO #											NO #	
OP1253	F	Ad-CRE	CD1	One small adenoma observed on tissue sections. Negative at µCT.																		YES #	
OP1255	F	Ad-CRE	CD1																			NO #	
OP1256	F	Ad-CRE	CD1																			NO #	
OP1257	F	Ad-CRE	CD1																			NO #	
OP1109	F	Ad-CRE	B6								NO #				NO #								NO #
OP1103	F	Ad-CRE	B6											NO #									NO #
					2days	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
EY3302	F	Ad-Cas9	B6		NO †																		
EY3341	F	Ad-Cas9	CD1		NO †																		
EY3314	F	Ad-Cas9	B6						NO #														
EY3342	F	Ad-Cas9	CD1						NO #														
EY3304	F	Ad-Cas9	B6						NO #		NO #												
EY3343	F	Ad-Cas9	CD1						NO #		NO #												
EY3313	F	Ad-Cas9	B6						NO #														
EY3303	F	Ad-Cas9	B6						NO #		NO #	NO †											
EY3306	F	Ad-Cas9	B6						NO #		NO #	NO †											
EY3308	F	Ad-Cas9	B6						NO #		NO #	NO †											
EY3310	F	Ad-Cas9	B6						NO #		NO #	NO †											
EY3311	F	Ad-Cas9	B6						NO #		NO #												
EY3344	F	Ad-Cas9	CD1						NO #		NO #				NO #								
EY3345	F	Ad-Cas9	CD1						NO #		NO #				NO #								
EY3346	F	Ad-Cas9	CD1						NO #		NO #				NO #								
EY3347	F	Ad-Cas9	CD1						NO #		NO #				NO #								
EY3348	F	Ad-Cas9	CD1						NO #		NO #				NO #								
EY3349	F	Ad-Cas9	CD1						NO #		NO #				NO #								
EY3350	F	Ad-Cas9	CD1						NO #		NO #				NO #								
					2days	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
OP1920	M	Ad-EA	B6		NO †																		
EY3316	F	Ad-EA	B6			NO †																	
OP1916	M	Ad-EA	B6			NO †																	
EY3339	F	Ad-EA	B6							YES †													
EY3340	F	Ad-EA	B6							YES †													
OP1119	F	Ad-EA	B6						YES †														
notag-2	F	Ad-EA	CD1						YES †														
notag-3	F	Ad-EA	CD1						YES †														
EY3318	F	Ad-EA	B6																				
EY3320	F	Ad-EA	B6								YES #												
EY3317	F	Ad-EA	B6								YES #												
EY3319	F	Ad-EA	B6								YES #												
EY3326	F	Ad-EA	B6								YES #												
EY3327	F	Ad-EA	B6								NO #												
EY3328	F	Ad-EA	B6								YES #				YES #								
EY3329	F	Ad-EA	B6								YES #				YES #								
EY3330	F	Ad-EA	B6								YES #				YES #								
OP1281	F	Ad-EA	CD1									YES #											
OP1277	F	Ad-EA	CD1									YES #											
OP1287	F	Ad-EA	CD1							NO #			YES #										
OP1288	F	Ad-EA	CD1										YES #										
OP1112	F	Ad-EA	B6								NO #			YES †									
OP1115	F	Ad-EA	B6										NO #		YES †								
OP1279	F	Ad-EA	CD1											YES #		YES †							
OP1252	F	Ad-EA	CD1											YES #		YES †							
OP1113	F	Ad-EA	B6								YES #			YES #		YES †							
OP1114	F	Ad-EA	B6								YES #			YES #		YES †							
OP1118	F	Ad-EA	B6								YES #			YES #		YES †							
OP1941	F	Ad-EA	B6								YES #			YES #		YES †							
OP1251	F	Ad-EA	CD1											YES #		YES †							
OP1259	F	Ad-EA	CD1	Vehicle Treatment (progression)										YES V #		YES V #							
OP1300	F	Ad-EA	CD1	Crizotinib Treatment (complete response)										YES C #		NO C #							
OP1258	F	Ad-EA	CD1	Crizotinib Treatment (complete response)										YES #		YES C #		NO C #					
OP1260	F	Ad-EA	CD1	Vehicle Treatment (progression)								YES #		YES #		YES V #		YES V #					
OP1290	F	Ad-EA	CD1	Crizotinib Treatment (complete response)										YES #		YES C #		NO C #					
OP1283	F	Ad-EA	CD1	Crizotinib Treatment (complete response)										YES #		YES C #		NO C #					
OP1298	F	Ad-EA	CD1	Crizotinib Treatment (complete response)										YES #		YES C #		YES C #		NO C #		NO C #	NO C #
OP1295	F	Ad-EA	CD1	Crizotinib Treatment (complete response)										YES #		YES C #		YES C #		NO C #		NO C #	NO C #
OP1260	F	Ad-EA	CD1											YES #		YES #		YES #				YES #	
OP1261	F	Ad-EA	CD1											YES #		YES #		YES #				YES #	
OP1262	F	Ad-EA	CD1									YES #		YES #		YES #		YES #				YES #	
OP1294	F	Ad-EA	CD1											YES #		YES #		YES #				YES #	
OP1296	F	Ad-EA	CD1											YES #		YES #		YES #				YES #	
OP1297	F	Ad-EA	CD1											YES #		YES #		YES #				YES #	
OP1299	F	Ad-EA	CD1											YES #		YES #		YES #				YES #	
OP1278	F	Ad-EA	CD1							NO #	YES #		YES #		YES #		YES #					YES #	
OP1289	F	Ad-EA	CD1											YES #		YES #		YES #				YES #	
OP1291	F	Ad-EA	CD1																			YES #	
OP1292	F	Ad-EA	CD1	Vehicle Treatment (progression)												YES #		YES V #		YES V #		YES V #	YES V #
OP1293	F	Ad-EA	CD1	Crizotinib Treatment (partial response)												YES #		YES C #		YES C #		YES C #	YES C #
OP1286	F	Ad-EA	CD1								YES #		YES #		YES #		YES #		YES #			YES #	
OP1942	F	Ad-EA	B6								YES #			YES #		YES #		YES #					YES †

Figure Legend

NO = No lung tumors detected
 YES = 1 or more lung tumors detected
 # = µCT scan
 † = necropsy & histology
 V = Vehicle (water)
 C = Crizotinib (100mg/kg/die)

Darker background color = Evidence for the presence of one or more lung lesions

This spreadsheet contains an annotated list of every mouse used in this study and the virus used for the intratracheal infection. The interval (in weeks) since infection is shown as a coloured horizontal bar. The time, outcome, and method of tumour detection are also reported. Symbols used are: YES = one or more tumour detected; NO = no tumours detected. # = evaluation by µCT; † = evaluation by necropsy and histopathology; V = mouse treated with vehicle (water); C = mouse treated with crizotinib (100 mg per kg per day).

Extended Data Table 2 | Response to crizotinib treatment

Mouse ID	Sex	Time (weeks) since infection at the start of treatment	Weeks treated	Treatment	Outcome at 2 weeks	Notes
OP1300	F	9.7	2	Crizotinib	Complete Response	Suppl. Videos 3 and 4
OP1290	F	12.3	2	Crizotinib	Complete Response	Suppl. Videos 7 and 8
OP1283	F	12.3	2	Crizotinib	Complete Response	
OP1258	F	11.0	2	Crizotinib	Complete Response	
OP1293	F	13.3	2	Crizotinib	Partial Response	Suppl. Videos 9 and 10
OP1295	F	12.0	2	Crizotinib	Complete Response	
OP1298	F	12.0	2	Crizotinib	Complete Response	
OP1280	F	11.0	2	Vehicle	Progression	Suppl. Videos 5 and 6
OP1259	F	12.0	2	Vehicle	Progression	Suppl. Videos 1 and 2
OP1292	F	13.3	2	Vehicle	Progression	

Table showing the response to crizotinib or vehicle treatment as judged by μ CT.

Extended Data Table 3 | Oligonucleotides used in this study

Name	Sequence
Alk_cDNA-rev	GGTCATGATGGTCGAGGTCC
Alk_Exon29_rev	GCTAGTGGAGTACAGGGCTC
Alk_gDNA-for (primer D Fig 1b and Supp Fig1b)	GCAGCGGGGCTTCCGAAGGGGC
Alk_gDNA-rev (primer C Fig1b and Supp Fig1b)	GTTTTACTGTGTCAGAAAGGG
Alk-rev	CAAGGCAGTGAGAACCTGAA
Eml4_cDNA-for	TGGAGTGGCAACTCACTAACAA
Eml4_cDNA-rev	GCAACTGCTCTAATGGTGCC
Eml4_Exon1_for	TAGAACTCGAGGCAAGATGGACGGTTTCGC
Eml4_gDNA-for (primer A Fig1b)	GCTCAAGAGGTGGGTTGTGT
Eml4_gDNA-rev (primer B Fig1b)	CAGGGCTGTGCCTAGATGAC
Eml4-for	GAGCCTTGTTGATACATCGTTC
Eml4-rev	TAGGAGGCAGTTTGGGCTAC
GAPDH_cDNA-for	ACCACAGTCCATGCCATCACTGCC
GAPDH_cDNA-rev	GTCTCGCTCCTGGAAGATGG
miR17-92_gDNA-for	TCGAGTATCTGACAATGTGG
miR17-92_gDNA-rev	TAGCCAGAAGTTCCAAATTGG
Npm1_cDNA-for	ACTACCTTTTCGGCTGTGAACT
Npm1_gDNA-for (primer A Supp Fig1b)	GTCTCTTGCCTCATTTGGGG
Npm1_gDNA-rev (primer B Supp Fig1b)	CTCCAGGAGCAGATCGCTTT

This table lists the names and sequences of each DNA oligonucleotide used in this study.

Extended Data Table 4 | Primer pairs and PCR reactions

Name	Description	Expected size (bp)
Alk_gDNA-for	Surveyor assay	961
Alk_gDNA-rev		
Eml4_gDNA-for	Surveyor assay	602
Eml4_gDNA-rev		
Eml4_gDNA-for	<i>Eml4-Alk</i> genomic	527
Alk_gDNA-rev		
Alk_gDNA-for	<i>Alk-Eml4</i> genomic	1036
Eml4_gDNA-rev		
Eml4_gDNA-for	Deletion	1044
Alk_gDNA-for		
miR17-92_gDNA-for	Control (gDNA)	255
miR17-92_gDNA-rev		
GAPDH_cDNA-for	Control (cDNA)	237
GAPDH_cDNA-rev		
Eml4-for	<i>Eml4-Alk</i> genomic (three primers)	<i>Eml4</i> : 240 <i>Eml4-Alk</i> : 190
Eml4-rev		
Alk-rev		
Eml4_cDNA-for	<i>Eml4-Alk</i> transcript (three primers)	<i>Eml4</i> : 336 <i>Eml4-Alk</i> : 276
Eml4_cDNA-rev		
Alk_cDNA-rev		
Eml4_cDNA-for	<i>Eml4-Alk</i> transcript (junction)	276
Alk_cDNA-rev		
Eml4_Exon1_for	<i>Eml4-Alk</i> transcript (full length)	3238
Alk_Exon29_rev		
Npm1_gDNA-for	<i>Npm1-Alk</i> genomic	581
Alk_gDNA-rev		
Alk_gDNA-for	<i>Alk-Npm1</i> genomic	1036
Npm1_gDNA-rev		
Npm1_cDNA-for	<i>Npm1-Alk</i> transcript	404
Alk_cDNA-rev		

This table lists the primer pairs and the sizes of the expected products for each PCR reaction described in this study.

Rapid modelling of cooperating genetic events in cancer through somatic genome editing

Francisco J. Sánchez-Rivera^{1,2*}, Thales Papagiannakopoulos^{1*}, Rodrigo Romero^{1,2}, Tuomas Tammela¹, Matthew R. Bauer¹, Arjun Bhutkar¹, Nikhil S. Joshi¹, Lakshmipriya Subbaraj¹, Roderick T. Bronson^{3,4}, Wen Xue¹ & Tyler Jacks^{1,2,5}

Cancer is a multistep process that involves mutations and other alterations in oncogenes and tumour suppressor genes¹. Genome sequencing studies have identified a large collection of genetic alterations that occur in human cancers^{2–4}. However, the determination of which mutations are causally related to tumorigenesis remains a major challenge. Here we describe a novel CRISPR/Cas9-based approach for rapid functional investigation of candidate genes in well-established autochthonous mouse models of cancer. Using a *Kras*^{G12D}-driven lung cancer model⁵, we performed functional characterization of a panel of tumour suppressor genes with known loss-of-function alterations in human lung cancer. Cre-dependent somatic activation of oncogenic *Kras*^{G12D} combined with CRISPR/Cas9-mediated genome editing of tumour suppressor genes resulted in lung adenocarcinomas with distinct histopathological and molecular features. This rapid somatic genome engineering approach enables functional characterization of putative cancer genes in the lung and other tissues using autochthonous mouse models. We anticipate that this approach can be used to systematically dissect the complex catalogue of mutations identified in cancer genome sequencing studies.

Lung cancer genome sequencing studies have revealed a multitude of recurrent mutations and copy number alterations^{2–4}. However, the determination of which mutations are causally related to tumorigenesis remains a major challenge. Genetically engineered mouse models of lung cancer have assisted in the functional characterization of putative driver events identified in human lung tumours^{6,7}, but these require modification of the germ line and cannot be performed in a highly parallel manner.

Recent work from our laboratory has demonstrated the feasibility of using the CRISPR (clustered regularly interspaced short palindromic repeats)/Cas9 system to directly mutate cancer genes in the liver following hydrodynamic delivery of plasmids carrying the CRISPR components⁸, which relies on the efficient transfection of hepatocytes. To rapidly interrogate cancer genes in the lung and other tissues, we developed pSECC (Fig. 1a), a lentiviral-based system that delivers both the CRISPR system and Cre recombinase. In this setting, CRISPR-induced mutation of genes can be examined in the context of several of the well-studied conditional Cre/*loxP* mouse models of lung cancer⁹ and other cancer types. To test this system, we used genetically engineered mouse models of lung adenocarcinoma, in which tumours are induced in *loxP*-Stop-*loxP* *Kras*^{G12D/+} (hereafter referred to as *Kras*^{LSL-G12D/+}) or *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} mice upon intratracheal administration of lentiviral vectors expressing Cre recombinase^{10,11}.

To validate pSECC, we developed the Green-Go (GG) reporter cell line, which expresses GFP following exposure to Cre (Extended Data Fig. 1a–c). To assess the efficiency of Cas9 in tumours *in vivo*, we targeted a Cre-activatable tdTomato knock-in reporter allele¹² with pSECC lentiviruses expressing a single guide RNA (sgRNA) against tdTomato (sgTom) or an empty vector control (Extended Data Fig. 1d, e). At

10 weeks post-infection, we assessed knockdown of tdTomato expression by immunohistochemistry. We observed that 28% of tumours lacked tdTomato expression, suggesting that the system was functional *in vivo* by editing an endogenous allele in the context of a lung tumour (Extended Data Fig. 2a–e). Importantly, animals infected with empty pSECC rarely contained non-tumour Tomato-expressing cells (data not shown), indicating that there is minimal infection of non-epithelial cells when using a low lentiviral titre.

We then proceeded to functionally characterize tumour suppressor genes using this approach. Loss of NK2 homeobox 1 (*Nkx2-1*), a master regulator of lung development¹³, or phosphatase and tensin homologue (*Pten*), a negative regulator of oncogenic PI(3)K/Akt signalling¹⁴ accelerates lung tumorigenesis in *Kras*^{LSL-G12D/+} and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} lung tumour models^{10,15,16}. We infected *Kras*^{LSL-G12D/+} and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} animals with pSECC vectors expressing validated sgPten, sgNkx2-1 and controls (sgTom and empty vector) to induce lung tumours. Ten weeks post-infection, we euthanized animals to assess the effects of CRISPR/Cas9-mediated gene editing in tumours by histopathology, survival assays and deep sequencing of the targeted alleles (Fig. 1a). All animals expressing sgRNAs targeting *Pten* or *Nkx2-1* contained tumours with marked histopathological differences compared to controls (Fig. 1b, d and Extended Data Fig. 3a–d).

Animals infected with sgNkx2-1-pSECC developed mucinous adenocarcinomas typified by the presence of elongated cells, mucin production and glandular rearrangements, in agreement with previous Cre/*loxP*-based (*Nkx2-1*^{fl/fl}) data¹⁶ (Fig. 1b). The majority of tumours (61%, 54/88 tumours) from sgNkx2-1-pSECC animals lacked *Nkx2-1* expression (compared to 0/33 tumours from controls) (Fig. 1b, c). Importantly, 85% (46/54 tumours) of these *Nkx2-1*-negative tumours stained positively for mucin (Fig. 1c), a biomarker of mucinous adenocarcinomas¹⁶. Thus, although a subset of tumours appeared to partially or fully escape CRISPR-mediated deletion of *Nkx2-1*, we were able to observe clear phenotypes by examining the full spectrum of tumours generated by sgNkx2-1-pSECC.

Animals infected with sgPten-pSECC demonstrated complete loss of Pten protein in 74% of tumours (40/54 tumours), which was accompanied by a concomitant increase in pAkt (S473), a downstream biomarker of increased PI(3)-kinase pathway activity (Fig. 1d, e). These results mimic previously published data using a *Pten*^{fl/fl} allele in *Kras*^{LSL-G12D/+} mice¹⁵. Collectively, these data indicate that CRISPR/Cas9-based gene editing leads to loss-of-function mutations in this model and closely parallels what is seen with the use of traditional conditional alleles.

We next used this system to study adenomatous polyposis coli (*Apc*), a tumour suppressor whose functional role in lung adenocarcinoma has not been characterized. Of note, *Apc* is found in a region that frequently undergoes copy number loss in human lung cancer⁴. We infected animals with pSECC lentiviruses expressing a validated sgRNA¹⁷ targeting

¹David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA. ²Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA. ³Tufts University, Boston, Massachusetts 02115, USA. ⁴Harvard Medical School, Boston, Massachusetts 02115, USA. ⁵Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

*These authors contributed equally to this work.

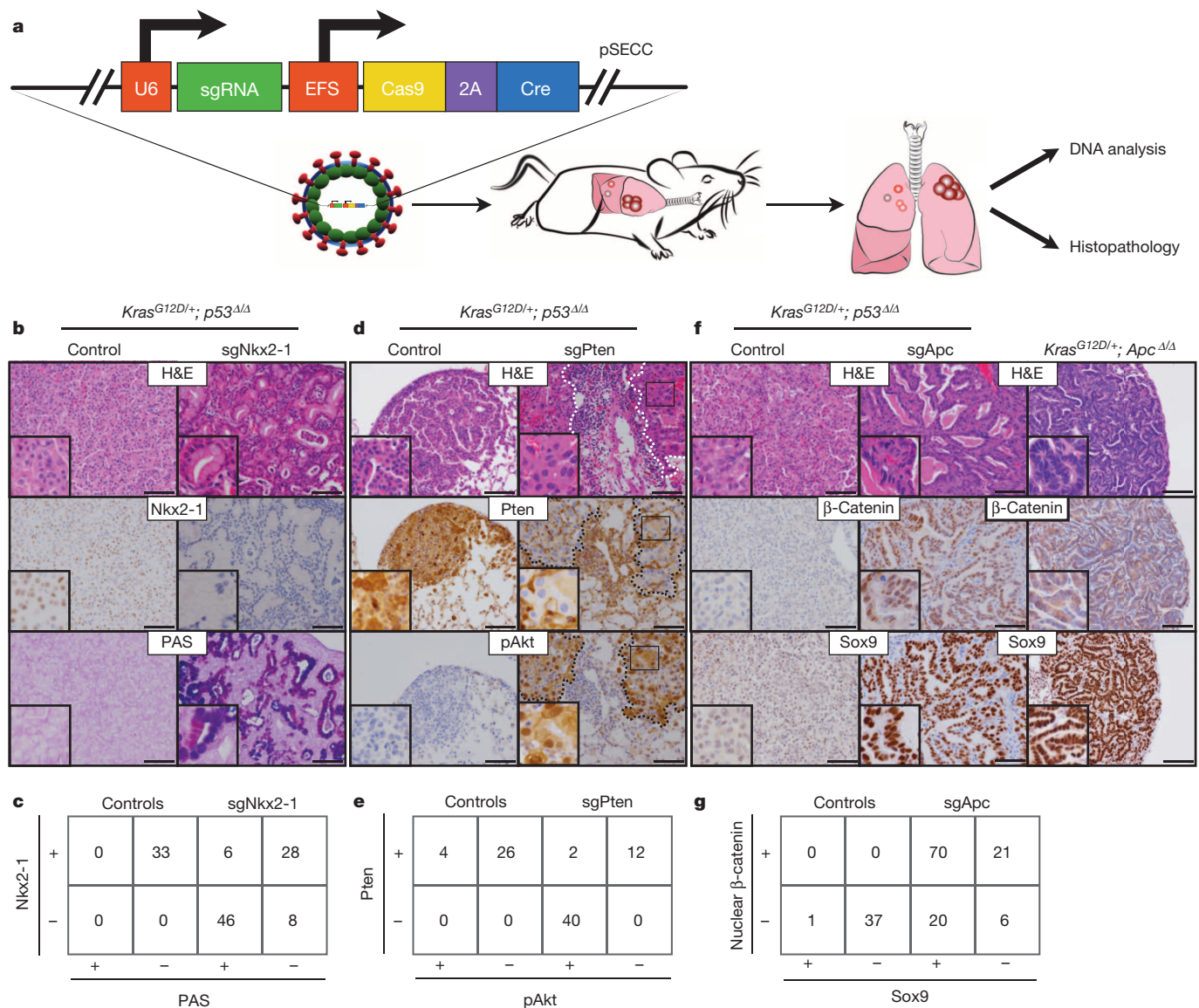


Figure 1 | CRISPR/Cas9-mediated somatic gene editing in an autochthonous mouse model of lung cancer. **a**, pSECC lentiviruses are intratracheally delivered into mouse lungs to delete genes of interest. DNA extracted from tumour-bearing lungs is analysed by high-throughput sequencing and surveyor assays to identify gene-editing events. The remaining tissue is analysed by histopathology. **b**, Representative haematoxylin and eosin (H&E) and immunohistochemistry (IHC) staining of serial sections from lung tumours of mice 10 weeks after infection with sgTom-pSECC (left panel) or sgNkx2-1-pSECC (right panel). Alcian Blue/PAS (periodic acid-Schiff) stain for mucin. Note the accumulation of mucin only in tumours from sgNkx2-1-pSECC mice. **c**, Contingency tables demonstrating anti-correlation between Nkx2-1 expression and mucin production (PAS stain) (two-sided Fisher's exact test, $P < 0.0001$). **d**, Representative H&E and IHC stainings of serial sections from lung tumours of mice 10 weeks after infection with

sgTom-pSECC (left panel) or sgPten-pSECC (right panel). Dashed lines demarcate tumour boundaries on each consecutive histological section. **e**, Contingency tables demonstrating anti-correlation between Pten expression and Akt phosphorylation (two-sided Fisher's exact test, $P < 0.0001$). **f**, Representative H&E and IHC stainings of serial sections from lung tumours of mice 10 weeks after infection with sgTom-pSECC (left panel) or sgApc-pSECC (middle panel). The far right panel corresponds to serial sections from lung tumours of *Kras^{LSL-G12D/+}; Apc^{fl/fl}* mice 18 weeks after infection with Adeno-Cre. **g**, Contingency tables demonstrating positive correlation between β-catenin expression and Sox9 expression (two-sided Fisher's exact test, $P < 0.0001$). These data are representative of at least 3 independent *Kras^{LSL-G12D/+}* or *Kras^{LSL-G12D/+}; p53^{fl/fl}* mice infected with each pSECC sgRNA. All scale bars, 0.05 mm.

Apc. At 10 weeks post-infection, we observed a striking difference in the histopathology of sgApc tumours compared to controls (Fig. 1f and Extended Data Fig. 3e). Importantly, tumours from *Kras^{LSL-G12D/+}; Apc^{fl/fl}* mice, which express a conditional allele of *Apc*¹⁸, exhibited identical histopathology (Fig. 1f). Tumours with Cas9-mediated deletion of *Apc* were highly dedifferentiated, invasive and had a significant stromal component (Fig. 1f). The majority of these tumours (78%, 91/117) stained strongly for nuclear β-catenin, a marker of *Apc* mutation in colon cancer and other settings¹⁹ (Fig. 1f, g). Furthermore, 77% (70/91) of tumours with nuclear β-catenin stained positive for the transcription factor Sox9, which might reflect a distal embryonic differentiation state^{20,21}. Of note,

we observed a statistically significantly higher number of Sox9-positive tumours in *Kras^{LSL-G12D/+}; p53^{fl/fl}*-sgApc (29/33, or 88%) than in *Kras^{LSL-G12D/+}*-sgApc mice (41/58 tumours, or 71%), suggesting a possible role for p53 in regulating this change in differentiation (Extended Data Fig. 6b, c).

To further characterize the differentiation state of sgApc tumours, we stained serial sections for lung differentiation markers, including Sox2, Clara cell secretory protein (CCSP), surfactant protein C (SP-C), p63, Nkx2-1 and Sox9 (Extended Data Fig. 6a)²². Tumours from *Kras^{LSL-G12D/+}; p53^{fl/fl}*-sgTom mice stained positively for CCSP, SP-C and Nkx2-1 and negatively for Sox2, p63 and Sox9. In contrast, tumours

from sgApc mice frequently stained positively for SP-C, Nkx2-1 and Sox9 and negatively for CCSP, Sox2 and p63. A large number of tumours from sgApc mice had areas with low levels or complete absence of Nkx2-1, which correlated with the levels of the Nkx2-1 transcriptional target SP-C²² (16/52 tumours, or 31%) (Extended Data Fig. 6d). These data indicate that these tumours are poorly differentiated and that hyperactivation of the canonical Wnt signalling pathway through loss of *Apc* in *Kras*-driven lung adenocarcinomas results in tumours with varying degrees of differentiation. These results also mimic what we observed in tumours from *Apc* conditional knockout mice (Fig. 1f and Extended Data Fig. 6e) and recapitulate recent observations in a *Braf*^{V600E}-driven mouse model of lung adenocarcinoma upon Wnt pathway hyperactivation²³.

Our initial analysis demonstrated histological and pathway-specific differences upon deletion of these tumour suppressors in lung tumours. To assess the overall impact of these alterations on tumorigenesis, we measured tumour burden and grade in both *Kras*^{LSL-G12D/+} and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} animals. Deletion of *Pten* and *Apc* significantly increased overall tumour burden, which correlated with higher tumour grades (Grade 3 and 4) (Fig. 2a–c and Extended Data Fig. 3f, g). *Nkx2-1* deletion had a significant effect on overall tumour burden only in *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} animals; however, we observed a striking transition to highly dedifferentiated mucinous adenocarcinoma tumours in both *Kras*^{LSL-G12D/+} and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} mice (Fig. 2a–c, Extended Data Fig. 3f, g and Extended Data Fig. 4a–d). Conversely, *Apc* deletion had a significant effect on tumour burden only in *Kras*^{LSL-G12D/+} mice (Fig. 2a and Extended Data Fig. 3f, g). Deletion of all three genes led to increased BrdU incorporation, suggesting that the increased tumour burden is partly due to increased proliferation (Extended Data Fig. 3h). These data demonstrate the tumour suppressive role of *Nkx2-1*, *Pten* and *Apc* in the context of oncogenic *Kras*. Furthermore, the unique histopathology observed for each targeted tumour suppressor gene in this *Kras*-driven model illustrates the potential of this approach to rapidly model the effects of cooperative genetic events in lung tumorigenesis and progression.

Using this *in vivo* somatic genome editing approach, we observed inter- and intra-tumoral heterogeneity in terms of CRISPR-based loss-of-function of *Pten* in sgPten animals (Fig. 2d, e and Extended Data Fig. 5). Clones that acquired loss of *Pten* had increased PI(3)K/Akt signalling and may, therefore, have had a selective advantage over tumours that retained wild-type *Pten* within the same animal. We observed that tumours with complete or sub-clonal loss of *Pten* were significantly larger than tumours that retained *Pten* (Fig. 2d, e).

The histopathological and immunohistochemistry analyses indicate that the pSECC system is highly efficient *in vivo*, leading to robust target-specific phenotypic differences in lung tumours. To confirm Cas9-mediated editing of the alleles and precisely characterize the events at single-nucleotide resolution, we performed deep sequencing of target loci from whole lung and tumour DNA. Within a 23 base pair (bp) window (± 10 bp flanking the protospacer adjacent motif (PAM) sequence at each locus), the rate of mutations observed in the sgTarget samples was significantly greater than in the control samples (Fig. 3a–c). Using the control samples as a background model to analyse the mutational rate revealed that sgTarget samples were enriched for mutations within 7 bp upstream of the PAM sequences in predicted cutting sites, strongly suggesting that they are not secondary consequences of tumour progression (Fig. 3d and Extended Data Fig. 7a–c). The maximum per-base mutation frequency observed in sgTarget samples was 71.7% in *Nkx2-1*, 66.06% in *Pten* and 39.91% in *Apc* (in contrast to control samples: 0.11%, 0.73% and 0.14%, respectively). On average, $27.48\% \pm 10.3$ (*Nkx2-1*), $44.64\% \pm 5.3$ (*Pten*) and $13.54\% \pm 5.3$ (*Apc*) read fragments covering this 7 bp locus harboured indels in sgTarget samples. Across all sgTarget samples, > 94% of observed indels constituted non-synonymous frame-altering events (Extended Data Fig. 7d, e and Supplementary Tables 3 and 4).

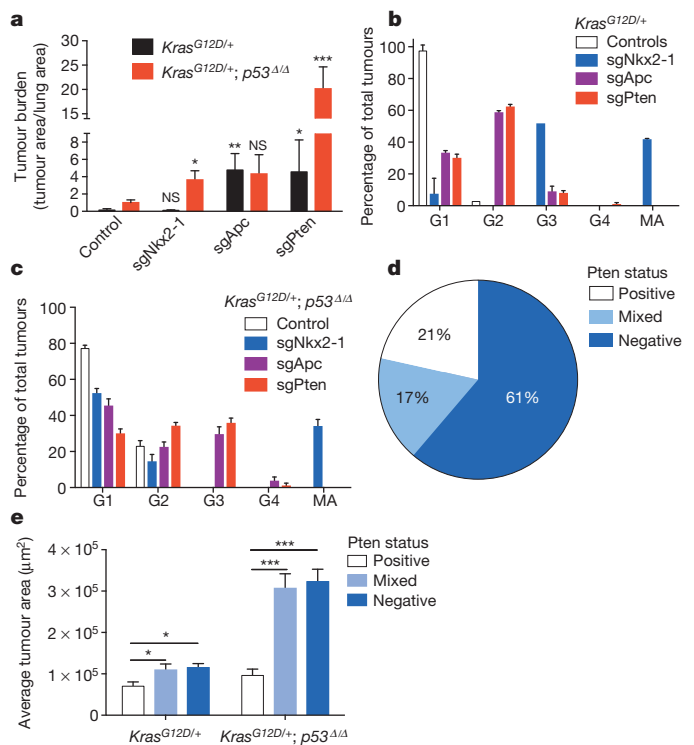


Figure 2 | Histopathological characterization of tumours from pSECC infected animals. **a**, Combined quantification of tumour burden (total tumour area/total lung area) in both *Kras*^{LSL-G12D/+} and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} animals 10 weeks after infection with pSECC lentiviruses expressing: control (empty or sgTom), *Kras*^{LSL-G12D/+} ($n = 4$) and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} ($n = 7$), sgNkx2-1 (*Kras*^{LSL-G12D/+} ($n = 2$) and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} ($n = 6$)), sgApc (*Kras*^{LSL-G12D/+} ($n = 3$) and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} ($n = 6$)) and sgPten (*Kras*^{LSL-G12D/+} ($n = 4$) and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} ($n = 3$)). The asterisks indicate statistical significance obtained from comparing *Kras*^{LSL-G12D/+} - sgTarget samples to *Kras*^{LSL-G12D/+} - control samples or *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} - sgTarget samples to *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} - control samples using Student's *t*-test (two-sided). **b**, **c**, Distribution of tumour grades in *Kras*^{LSL-G12D/+} (**b**) or *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} (**c**) animals 10 weeks after infection with pSECC lentiviruses expressing: control (empty or sgTom), *Kras*^{LSL-G12D/+} ($n = 4$) and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} ($n = 7$), sgNkx2-1 (*Kras*^{LSL-G12D/+} ($n = 2$) and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} ($n = 6$)), sgApc (*Kras*^{LSL-G12D/+} ($n = 3$) and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} ($n = 6$)) and sgPten (*Kras*^{LSL-G12D/+} ($n = 4$) and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} ($n = 3$)). G1, grade 1; G2, grade 2; G3, grade 3; G4, grade 4; MA, mucinous adenocarcinoma. **d**, Distribution of Pten IHC staining status in all sgPten-pSECC infected animals ($n = 9$) represented as percent of negative, mixed and positive tumours. **e**, Quantification of average tumour area (μm^2) of tumours staining negative, mixed or positive in all sgPten-pSECC infected animals ($n = 9$). Positive tumour, ~100% of the tumour cells stained positive for Pten. Mixed tumour, at least ~30% of tumour cells stained positive for Pten. Negative tumour, < 25% of the tumour cells stained positive for Pten. NS, not significant, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ obtained from two-sided Student's *t*-test. All error bars denote s.e.m.

Several studies have reported that Cas9 can bind to sites in the genome other than the intended target site^{24–27}, which could result in unintended editing at an off-target (OT) site. To assess off-target editing, we analysed the top three predicted²⁴ loci (Supplementary Tables 2) for each sgRNA by deep sequencing. We observed negligible off-target editing (Extended Data Fig. 8). On average, $0.048\% \pm 0.031$ (for sgNkx2-1), $0.26\% \pm 0.096$ (for sgPten) and $0.051\% \pm 0.027$ (for sgApc) of read fragments harboured indels in the off-target sites (Supplementary Tables 5–7). This data suggests that the observations reported for each of the sgRNAs arise from deletion of the intended target and not from editing of another gene.

The goal of cancer genomics is to identify genetic events that underlie cancer initiation and progression. The functional interrogation of

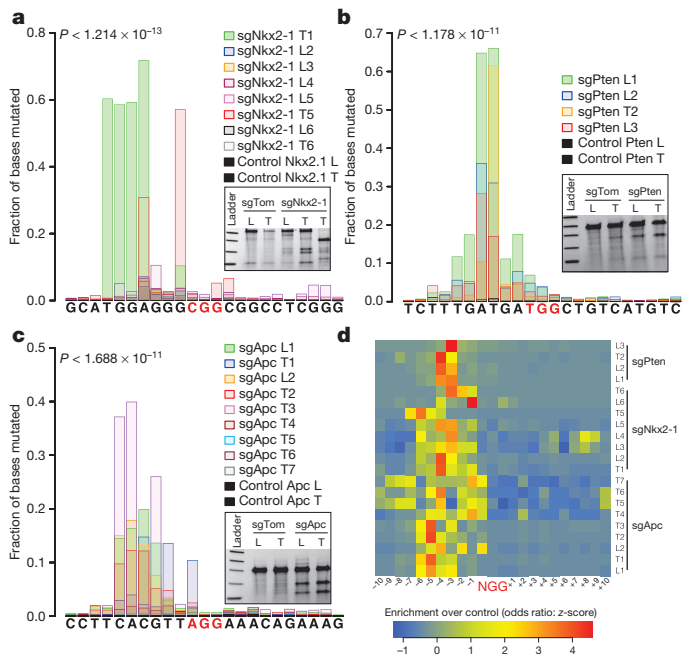


Figure 3 | CRISPR/Cas9 efficiently generates insertions and deletions (indels) in autochthonous tumours. **a–c**, Fraction of bases mutated per position in 10 bp flanks on either side of the protospacer adjacent motif (PAM) sequence (highlighted in red). Samples were obtained from entire lobes (L) or microdissected tumours (T) from mice 10 weeks after infection with pSECC lentiviruses targeting *Nkx2-1* (**a**), *Pten* (**b**) or *Apc* (**c**). *P* values denote enrichment of mutation rate in sgTarget-pSECC samples compared to sgTom-pSECC control samples (Wilcoxon rank sum test). Insets depict surveyor assays for each of the targets from either entire lobes (L) or microdissected tumours (T) from mice. Samples obtained from mice infected with sgTom-pSECC were used as controls. **d**, Positional enrichment of mutations in sgTarget-pSECC samples compared to sgTom-pSECC control samples based on all mutations considered at a given position (SNPs, indels). Each row represents a different sgRNA lung (L) or tumour (T) sample. Each cell represents the row-normalized (*z*-score) odds ratio estimate of mutational enrichment over an associated control sample (Fisher's exact test) upstream (+) or downstream (–) of the PAM sequence.

putative cancer genes in appropriate experimental models will elucidate which mutations identify bona fide cancer genes. This study presents a novel approach to rapidly evaluate human cancer genome candidates and assess cooperativity between genetic events in the context of well-established mouse models of lung cancer. Moreover, our ability to model different lung adenocarcinoma subtypes allows for the detailed study of subtype-specific molecular mechanisms controlling disease initiation and progression. We anticipate that this approach can be readily adapted to many existing Cre/*loxP*-based genetically engineered mouse models of several cancer types to facilitate the rapid functional assessment of new hypotheses generated by cancer genome studies.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 28 July; accepted 2 October 2014.

Published online 22 October 2014.

- Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).
- Imielinski, M. *et al.* Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**, 1107–1120 (2012).
- Govindan, R. *et al.* Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* **150**, 1121–1134 (2012).

- The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543–550 (2014).
- Jackson, E. L. *et al.* Analysis of lung tumor initiation and progression using conditional expression of oncogenic *K-ras*. *Genes Dev.* **15**, 3243–3248 (2001).
- McFadden, D. G. *et al.* Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. *Cell* **156**, 1298–1311 (2014).
- Frese, K. K. & Tuveson, D. A. Maximizing mouse cancer models. *Nature Rev. Cancer* **7**, 645–658 (2007).
- Xue, W. *et al.* CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature* **514**, 380–384 (2014).
- Farago, A. F., Snyder, E. L. & Jacks, T. SnapShot: Lung cancer models. *Cell* **149**, 246–246.e1 (2012).
- Winslow, M. M. *et al.* Suppression of lung adenocarcinoma progression by *Nkx2-1*. *Nature* **473**, 101–104 (2011).
- DuPage, M. *et al.* Endogenous T cell responses to antigens expressed in lung adenocarcinomas delay malignant tumor progression. *Cancer Cell* **19**, 72–85 (2011).
- Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neurosci.* **13**, 133–140 (2010).
- Rock, J. R. & Hogan, B. L. Epithelial progenitor cells in lung development, maintenance, repair, and disease. *Annu. Rev. Cell Dev. Biol.* **27**, 493–512 (2011).
- Song, M. S., Salmena, L. & Pandolfi, P. P. The functions and regulation of the PTEN tumour suppressor. *Nature Rev. Mol. Cell Biol.* **13**, 283–296 (2012).
- Curry, N. L. *et al.* *Pten*-null tumors cohabiting the same lung display differential AKT activation and sensitivity to dietary restriction. *Cancer Discov.* **3**, 908–921 (2013).
- Snyder, E. L. *et al.* *Nkx2-1* represses a latent gastric differentiation program in lung adenocarcinoma. *Mol. Cell* **50**, 185–199 (2013).
- Schwank, G. *et al.* Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* **13**, 653–658 (2013).
- Cheung, A. F. *et al.* Complete deletion of *Apc* results in severe polyposis in mice. *Oncogene* **29**, 1857–1864 (2010).
- Moon, R. T., Kohn, A. D., De Ferrari, G. V. & Kaykas, A. WNT and β -catenin signalling: diseases and therapies. *Nature Rev. Genet.* **5**, 691–701 (2004).
- Pacheco-Pinedo, E. C. *et al.* Wnt/ β -catenin signaling accelerates mouse lung tumorigenesis by imposing an embryonic distal progenitor phenotype on lung epithelium. *J. Clin. Invest.* **121**, 1935–1945 (2011).
- Kornish, J. D., Sinner, D. & Zorn, A. M. Interactions between SOX factors and Wnt/ β -catenin signaling in development and disease. *Dev. Dyn.* **239**, 56–68 (2010).
- Hogan, B. L. *et al.* Repair and regeneration of the respiratory system: complexity, plasticity, and mechanisms of lung stem cell function. *Cell Stem Cell* **15**, 123–138 (2014).
- Juan, J., Muraguchi, T., Iezza, G., Sears, R. C. & McMahon, M. Diminished WNT \rightarrow β -catenin \rightarrow c-MYC signaling is a barrier for malignant progression of BRAF^{V600E}-induced lung tumors. *Genes Dev.* **28**, 561–575 (2014).
- Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnol.* **31**, 827–832 (2013).
- Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnol.* **31**, 822–826 (2013).
- Wu, X. *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nature Biotechnol.* **32**, 670–676 (2014).
- Kuscu, C., Arslan, S., Singh, R., Thorpe, J. & Adli, M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nature Biotechnol.* **32**, 677–683 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. McFadden and Y. Soto-Feliciano for critical reading of the manuscript, H. Yin, S. Levine and T. Mason for MiSeq sequencing support, R. Stott, J. Bartlebaugh and C. Shivalila for technical assistance and K. Cormier and C. Condon from the Hope Babette Tang (1983) Histology Facility for technical support. This work was supported by the Howard Hughes Medical Institute, the Ludwig Center for Molecular Oncology at MIT and in part by Cancer Center Support (core) grant P30-CA14051 from the National Cancer Institute. T.P. is supported by the Hope Funds for Cancer Research. T.J. is a Howard Hughes Medical Institute Investigator, the David H. Koch Professor of Biology, and a Daniel K. Ludwig Scholar.

Author Contributions F.J.S.-R, T.P. and T.J. designed the study; F.J.S.-R, T.P., R.R., M.R.B. and L.S. performed experiments; T.T. generated *Kras*^{LSL-G12D/+}; *Apc*^{fl/fl} data; A.B. conducted bioinformatic analyses; N.S.J. generated GG cells; R.T.B. provided pathology assistance; W.X. gave conceptual advice; F.J.S.-R, T.P. and T.J. wrote the manuscript with comments from all authors.

Author Information Illumina MiSeq sequence datasets have been deposited into the NCBI repository under BioProjectID PRJNA256245. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.J. (tjacks@mit.edu).

METHODS

Lentiviral vectors and sgRNA cloning. The U6-sgRNA-EFS-Cas9-2A-Cre (pSECC) lentiviral vector was constructed by assembling four parts with overlapping DNA ends using Gibson assembly. Briefly, a 2.2 kb part (corresponding to the U6-Filler fragment from LentiCRISPR²⁸), a 0.3 kb part (corresponding to the EFS promoter from LentiCRISPR²⁸), a 5.3 kb part (corresponding to a Cas9-2A-Cre fragment, which was generated by assembly PCR) and a 5.7 kb lentiviral backbone were assembled using Gibson assembly following manufacturer guidelines. Detailed cloning strategies and primer sequences are available on request. For sgRNA cloning, the pSECC vector was digested with BsmBI and ligated with BsmBI-compatible annealed oligos (Supplementary Table 1). sgRNAs were designed using CRISPR Design²⁴ (which was also used to predict potential off-target sites; see Extended Data Fig. 8 and Supplementary Table 2) or E-CRISPR²⁹, except for sgApc which was previously reported¹⁷. An extra G (required for U6 transcriptional initiation) was added to the 5' end of sgRNAs that lacked it. The pSECC lentiviral vector is available through Addgene.

Lentiviral production. Lentiviruses were produced by co-transfection of 293T cells with lentiviral backbone constructs and packaging vectors (delta8.2 and VSV-G) using TransIT-LT1 (Mirus Bio). Supernatant was collected 48 and 72 h post-transfection, concentrated by ultracentrifugation at 25,000 r.p.m. for 90 min and resuspended in an appropriate volume of OptiMEM (Gibco).

Cell culture and generation of Green-Go cells. Cells were maintained in DMEM supplemented with 10% fetal bovine serum and gentamicin. Green-Go cells were generated by transducing 3T3 cells³⁰ with a bicistronic retrovirus containing an LTR promoter-driven inverted GFP (flanked by two sets of incompatible *loxP* sites) and a PGK-driven puromycin resistance cassette. Transduced cells were selected with puromycin and a single cell clone that expressed high levels of GFP 2–3 days after infection with a lentivirus expressing Cre recombinase was chosen.

Immunoblotting. Cells were lysed with ice-cold RIPA buffer (Pierce, #89900) supplemented with 1 × Complete Mini inhibitor mixture (Roche, #11 836 153 001) and mixed on a rotator at 4 °C for 30 min. Protein concentration of the cell lysates was quantified using the Bio-Rad DC Protein Assay (Catalogue #500-0114). Then 50–80 µg of total protein was separated on 4–12% Bis-Tris gradient gels (Life Technologies) by SDS-PAGE and then transferred to nitrocellulose membranes. The following antibodies were used for immunoblotting: anti-Flag (Sigma, F1804, 1:1,000), anti-Hsp90 (BD, #610418, 1:10,000), anti-Pten (Cell Signaling, 9188, 1:1,000), anti-TTF1 / Nkx2-1 (Epitomics, EP1584Y, 1:1,000).

Mice. All animal studies described in this study were approved by the MIT Institutional Animal Care and Use Committee. All animals were maintained on a mixed C57BL/6J × 129SvJ genetic background. *Kras*^{LSL-G12D/+} and *p53*^{fl/fl} mice have already been described^{5,31}. Mice were infected intratracheally with lentiviruses as described³². We infected a total of 7 mice with Empty-pSECC and 6 mice with sgTom-pSECC (for a total of 13 control mice), as well as 8 mice with sgNkx2-1-pSECC, 9 mice with sgPten-pSECC and 9 mice with sgApc-pSECC. No randomization or blinding was used. Total lung area occupied by tumour was measured on haematoxylin and eosin (H&E) stained slides using NIS-elements software.

Immunohistochemistry. Mice were euthanized by carbon dioxide asphyxiation. Lungs were perfused through the trachea with 4% paraformaldehyde (PFA), fixed overnight, transferred to 70% ethanol and subsequently embedded in paraffin. Sections were cut at a thickness of 4 µm and stained with H&E for pathological examination. Immunohistochemistry (IHC) was performed on a Thermo Autostainer 360 machine. Slides were antigen retrieved using Thermo citrate buffer, pH 6.0 in the pre-treatment module. Sections were treated with Biocare rodent block, primary antibody, and anti-mouse (Biocare) or anti-rabbit (Vector Labs) HRP-polymer. The slides were developed with Thermo Ultra DAB and counterstained with haematoxylin in a Thermo Gemini stainer and coverslips added using the Thermo Consul cover slipper. The following antibodies were used for IHC: anti-TTF1 / Nkx2-1 (Epitomics, EP1584Y, 1:1,200), anti-Pten (Cell Signaling, 9559, 1:100), anti-pAkt S473 (Cell Signaling, 4060, 1:100), anti-BrdU (Abcam, 6326, 1:100), anti-β-catenin (BD, 610154, 1:100), anti-Sox9 (Millipore, AB5535, 1:500), anti-RFP (Rockland, 600-401-379, 1:400), anti-Sox2 (Cell Signaling, 3728, 1:250), anti-CCSP (Millipore, 07-623, 1:2,000), anti-SP-C (Chemicon, AB3786, 1:1,000) and anti-p63 (Neomarkers, MS-1081, 1:200). To detect mucin, sections were stained with 1% Alcian Blue pH 2.5 and periodic acid-Schiff reagent. All pictures were obtained using a Nikon 80i microscope with a DS-U3 camera and NIS-elements software.

Genomic DNA isolation and Surveyor assay. Genomic DNA from entire snap-frozen left lung lobes or microdissected tumours was isolated using the High Pure PCR Template Preparation Kit (Roche) following manufacturer guidelines. PCR products for surveyor assay were amplified using Herculase II Fusion DNA polymerase (Agilent) (see Supplementary Table 1 for primers used for surveyor assay), gel purified and subsequently assayed with the Surveyor Mutation Detection Kit (Transgenomic). DNA was separated on 4–20% Novex TBE Gels (Life Technologies) and stained with ethidium bromide.

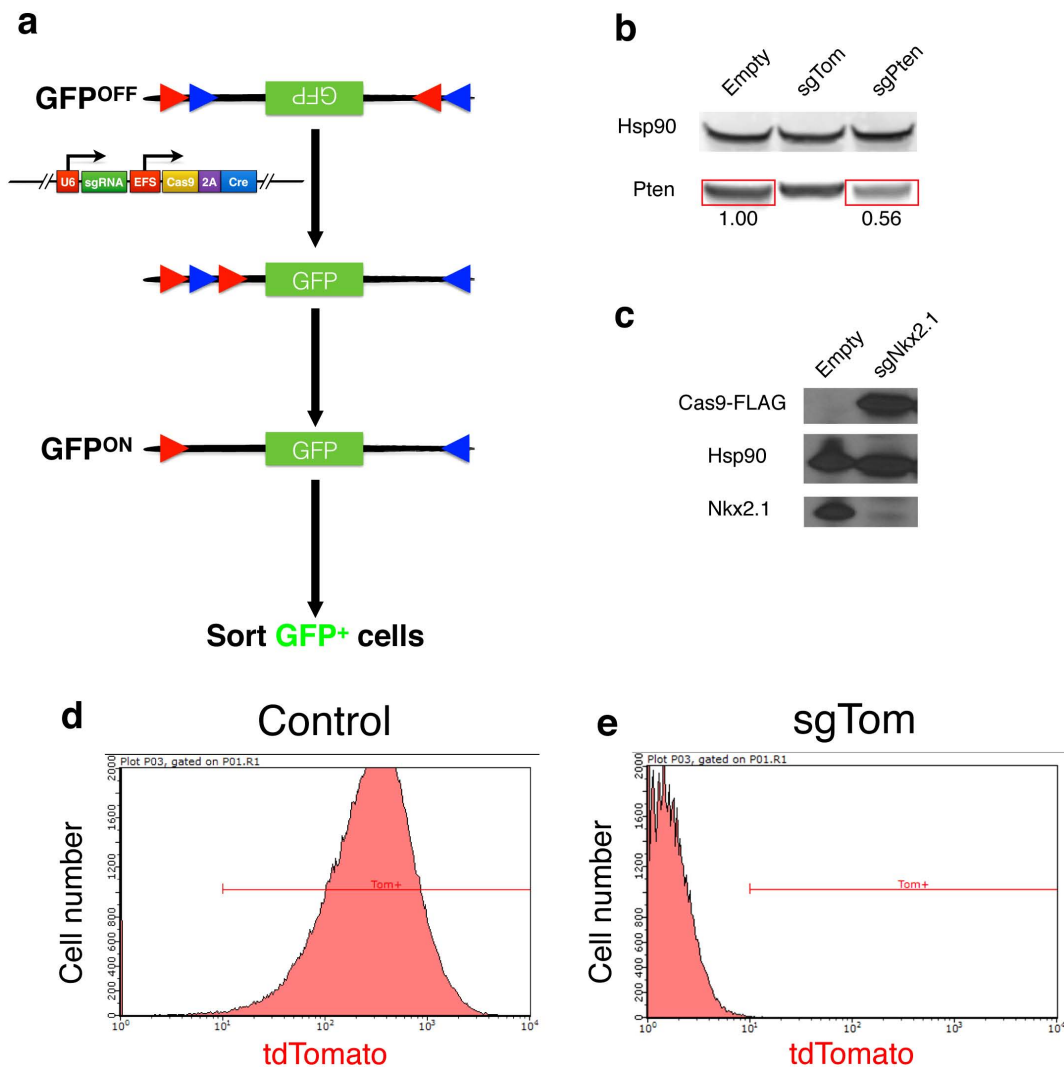
Deep sequencing and bioinformatic analysis of Cas9 target loci. For each target gene or potential off-target site, a genomic region containing the target sequence was amplified using Herculase II Fusion DNA polymerase and gel purified (primer sequences are shown in Supplementary Table 1). Sequencing libraries were prepared from 50 ng of PCR product using the Nextera DNA Sample Preparation Kit (Illumina) and sequenced on Illumina MiSeq machines. In order to retain high-quality sequence for mutation analysis, Illumina MiSeq reads (150 bp paired-end) were trimmed to 100mer paired end reads to drop lower quality 3' ends of reads. Traces of Nextera adapters were clipped from PE1 and PE2 100mer reads using the FASTX toolkit (Hannon Lab, CSHL). Reads greater than 15 nucleotides in length were retained. Additionally, reads with 50% or more bases below a base quality threshold of Q30 were dropped from subsequent analysis. Reference sequences with 10 bp genomic flanks were indexed using the Burrows–Wheeler Aligner (BWA) IS linear time algorithm³³ and reads were aligned using the BWA aligner. Reads with mapping quality greater than zero were retained. Overlapping alignments of paired end reads due to short inserts were resolved in order to avoid double counting of coverage and/or mutations observed in a single fragment. In order to minimize alignment ambiguity in the presence of mutations (including indels), the GATK Toolkit³⁴ was used to realign pooled cohorts mapping to a given locus. Mutations (base substitutions, insertions and deletions) were assessed using a combination of Samtools³⁵ and Annovar³⁶ (indel quantification and annotation), NGSUtils/BAMutils software suite³⁷ (total mutations per position), and custom scripts. Mutation frequencies were adjusted for sample purity (see next section) and per base substitution, insertion, and deletion frequencies were determined. Significance of overall mutation rates across 10 bp flanking the target locus was assessed using the Wilcoxon rank sum test comparing control and sgTarget sample events. Positional enrichment for mutation frequency compared to control samples was assessed using the conditional maximum likelihood odds ratio estimate (Fisher's exact test) and was mean centred and scaled (z-scores) across a 10 bp flank on either side of the PAM sequence in each sample. A number of other utilities/tools were used to enable various parts of the analysis, including: BEDTools³⁸, the Integrated Genome Viewer (IGV)³⁹, and Picard (<http://broadinstitute.github.io/picard/>). Statistical analyses and sequence enrichment plots were implemented in R (<http://www.R-project.org/>). Illumina MiSeq sequence data sets have been deposited into the NCBI repository under BioProjectID PRJNA256245.

Tumour purity correction. Lung lobe and microdissected tumour genomic DNA was used to perform real-time PCR based analysis to detect the relative levels of the un-recombined *Kras*^{LSL-G12D} allele (from non-tumour tissue) using forward primer: 5'-CTCTTGCCACGCCACCAGCTC-3' and reverse primer: 5'-AGCTAGCCACCATGGCTTGAGTAAGTCTGCA-3'. To correct for DNA loading of each sample, we amplified the chr5 10054507–10054621 region using forward primer: 5'-GAAGAAATAGAGGGCATGCTTC-3' and reverse primer: 5'-CTTCTCCAGTGACCTTATGTA-3'. Real-time PCR reactions were performed using KAPA Fast SYBR master mix in a Roche LightCycler Real-Time PCR instrument. To calculate percent purity we performed the following calculations for each sample: $\Delta C_p^{\text{tumourX}} = C_p^{\text{Chr5}} - C_p^{\text{KrasLSL-G12D/+}}$ to normalize for sample loading followed by $1/\Delta \Delta C_p = 1/(\Delta C_p^{\text{tumourX}} - \Delta C_p^{\text{LungControl}})$ for each sample to compare relative purity to lung tissue from *Kras*^{LSL-G12D/+} animals that were not infected with Cre. To validate the assay, we generated mouse embryonic fibroblasts from *Kras*^{LSL-G12D/+} mice treated with Cre recombinase (or control FipO recombinase). Purity values are reported in Supplementary Table 3.

Statistics. P values were determined by Student's *t*-test for all measurements of tumour burden and IHC quantifications except for contingency tables, in which Fisher's exact test or Chi-square test were used. All error bars denote s.e.m.

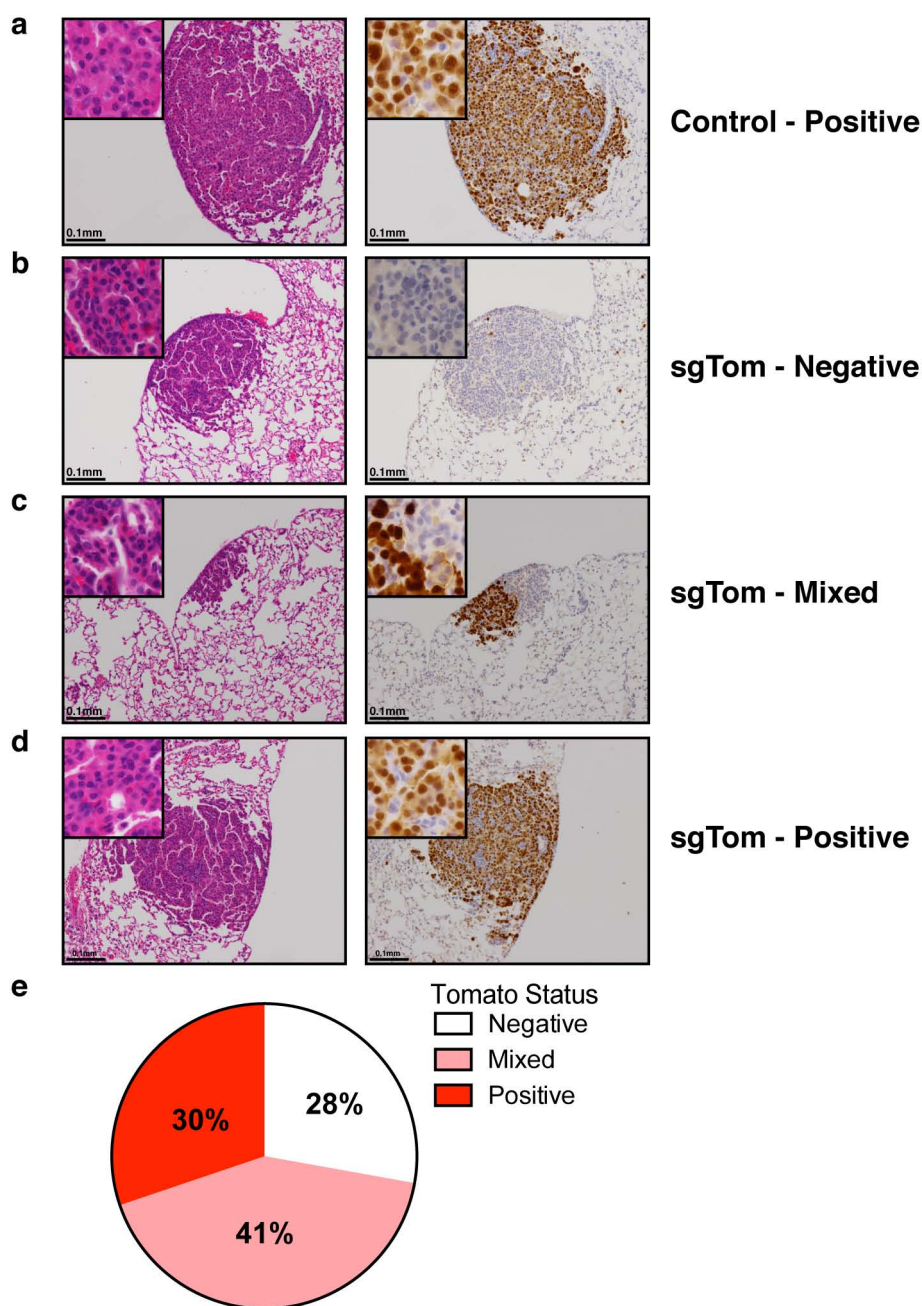
- Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014).
- Heigwer, F., Kerr, G. & Boutros, M. E-CRISP: fast CRISPR target site identification. *Nature Methods* **11**, 122–123 (2014).
- Psarras, S. *et al.* Gene transfer and genetic modification of embryonic stem cells by Cre- and Cre-PR-expressing MESV-based retroviral vectors. *J. Gene Med.* **6**, 32–42 (2004).
- Jackson, E. L. *et al.* The differential effects of mutant p53 alleles on advanced murine lung cancer. *Cancer Res.* **65**, 10280–10288 (2005).
- DuPage, M., Dooley, A. L. & Jacks, T. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nature Protocols* **4**, 1064–1072 (2009).
- Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
- McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
- Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).

37. Breese, M. R. & Liu, Y. NGSUtils: a software suite for analyzing and manipulating next-generation sequencing datasets. *Bioinformatics* **29**, 494–496 (2013).
38. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
39. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).



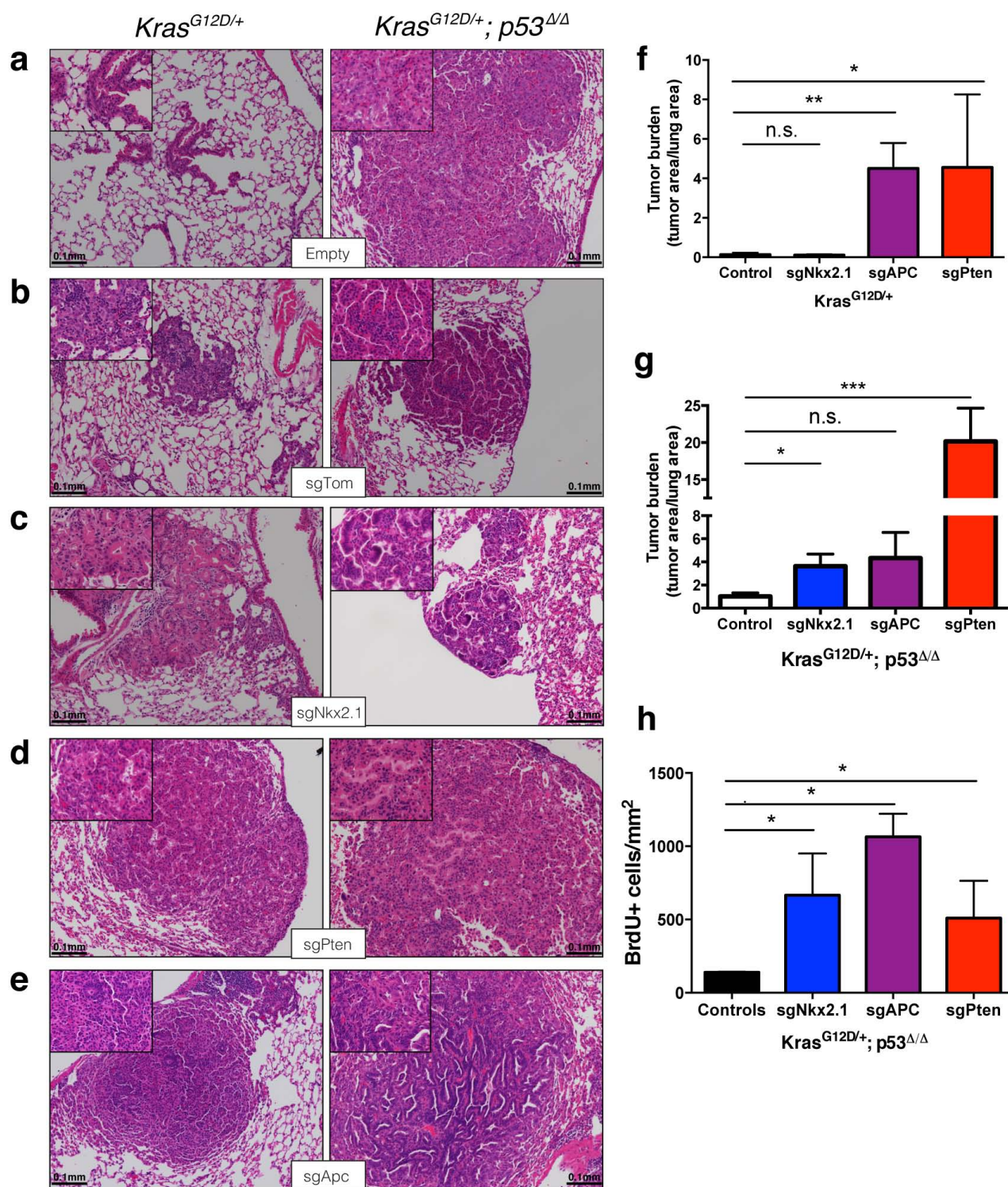
Extended Data Figure 1 | In vitro validation of pSECC. **a**, The Green-Go Cre-reporter cell line used to validate pSECC lentiviruses *in vitro*. Upon infection with a Cre-containing lentivirus, such as pSECC, cells become GFP⁺, allowing for purification of pSECC-containing cells by FACS. Red and blue triangles denote pairs of *loxP* sites, with red *loxP* sites being able to recombine only with other red *loxP* sites and blue *loxP* sites being able to recombine only with other blue *loxP* sites. **b**, Validation of sgPten-pSECC. Numbers below

the bands denote quantitation of protein level relative to empty vector control. **c**, Validation of sgNkx2-1-pSECC in a cell line that expresses Nkx2-1. **d**, **e**, Validation of sgTom-pSECC by fluorescence activated cell sorting (FACS). Briefly, a cell line obtained from a *Kras*^{LSL-G12D/+}; *p53*^{fl/fl}; *Rosa26*^{LSL-tdTomato/LSL-tdTomato} mouse was infected with either empty-pSECC (**d**) or sgTom-pSECC (**e**) and cultured for 10 days post-infection, after which time the cells were collected and analysed by FACS.



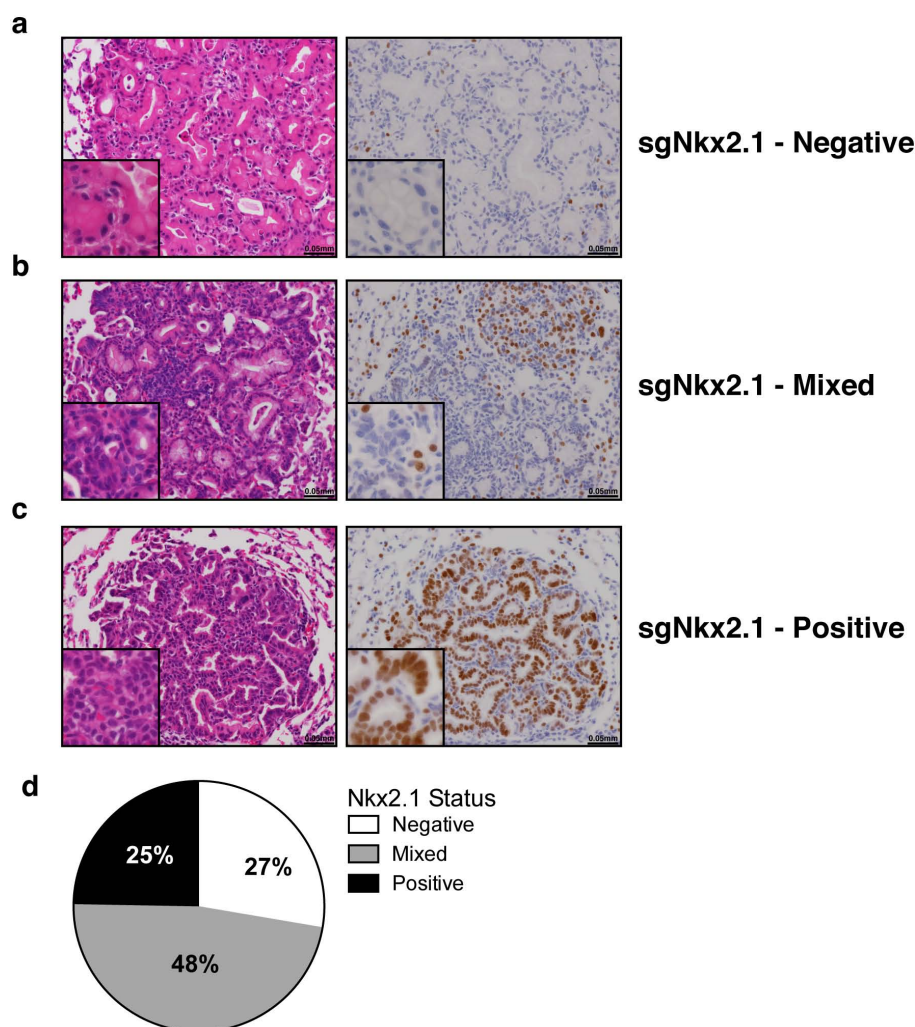
Extended Data Figure 2 | In vivo validation of pSECC. **a**, Representative H&E and tdTomato IHC staining of serial sections from lung tumours of $Kras^{LSL-G12D/+}; p53^{fl/fl}; Rosa26^{LSL-tdTomato/LSL-tdTomato}$ mice infected with Empty-pSECC. **b–d**, Representative H&E and IHC staining of serial sections from negative (**b**), mixed (**c**) and positive (**d**) lung tumours of $Kras^{LSL-G12D/+};$

$p53^{fl/fl}; Rosa26^{LSL-tdTomato/LSL-tdTomato}$ mice infected with sgTom-pSECC ($n = 6$). **e**, Distribution of lung tumours from all mice infected with sgTom-pSECC ($n = 6$) that were scored as negative, mixed or positive based on tdTomato IHC.



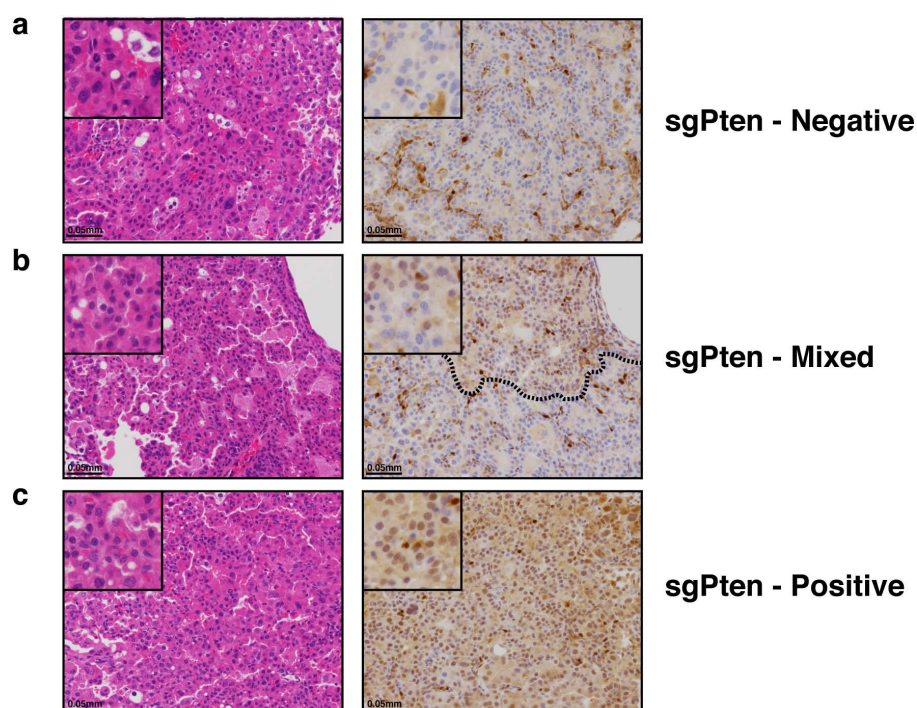
Extended Data Figure 3 | Histological analysis of lung tumours obtained from mice infected with pSECC lentiviruses. a–e, Representative H&E images of lung tumours obtained from mice infected with Empty-pSECC (a), sgTom-pSECC (b), sgNkx2-1-pSECC (c), sgPten-pSECC (d), and sgApc-pSECC (e). f, g, Quantification of tumour burden (total tumour area/total lung area) in *Kras*^{LSL-G12D/+} (f) or *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} (g) animals 10 weeks after infection with pSECC lentiviruses expressing: control (empty or sgTom, *Kras*^{LSL-G12D/+} (*n* = 4) and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} (*n* = 7)), sgNkx2-1 (*Kras*^{LSL-G12D/+} (*n* = 2) and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} (*n* = 6)), sgApc

(*Kras*^{LSL-G12D/+} (*n* = 3) and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} (*n* = 6)) and sgPten (*Kras*^{LSL-G12D/+} (*n* = 4) and *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} (*n* = 3)). h, Quantification of BrdU incorporation (BrdU⁺ cells per mm²) to assess proliferation of tumour cells from lung tumours in *Kras*^{LSL-G12D/+}; *p53*^{fl/fl} animals 10 weeks after infection with pSECC lentiviruses expressing: control (empty or sgTom, *n* = 4 tumours), sgNkx2-1 (*n* = 11 tumours), sgApc (*n* = 10 tumours) and sgPten (*n* = 15 tumours). Mice were given a pulse of BrdU for 4 h before being euthanized. n.s., not significant, **P* < 0.05, ***P* < 0.01, ****P* < 0.001 obtained from two-sided Student's *t*-test. All error bars denote s.e.m.



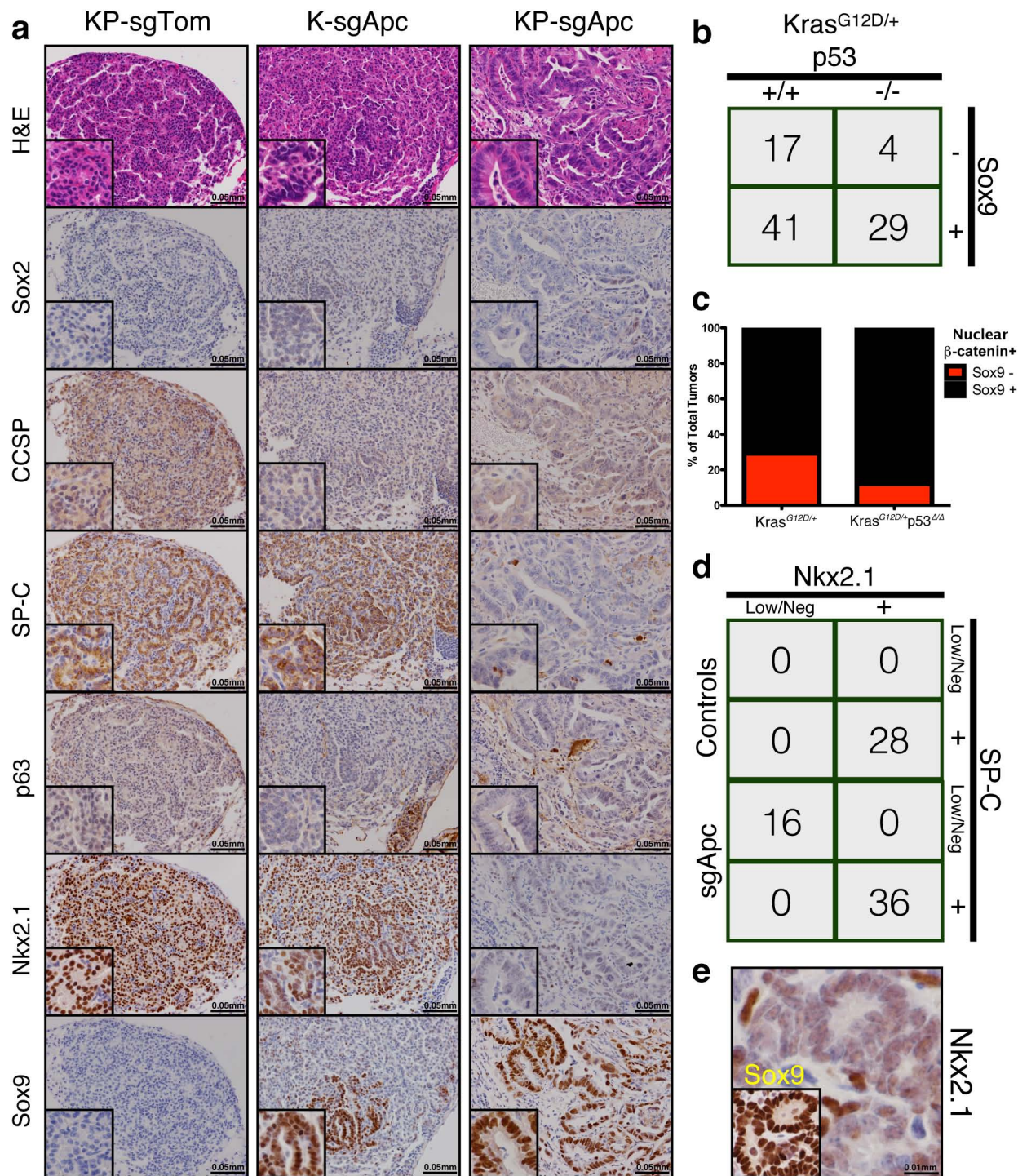
Extended Data Figure 4 | IHC-based analysis of mice infected with sgNkx2-1-pSECC. **a–c**, Negative (**a**), mixed (**b**) and positive (**c**) lung tumours of mice infected with sgNkx2-1-pSECC. **d**, Distribution of Nkx2-1 IHC staining status in all sgNkx2-1-pSECC infected animals ($n = 8$) represented as

percent of negative, mixed and positive tumours. Positive tumour, ~100% of the tumour cells stained positive for Nkx2-1. Mixed tumour, at least ~30% of tumour cells stained positive for Nkx2-1. Negative tumour, < 25% of the tumour cells stained positive for Nkx2-1.



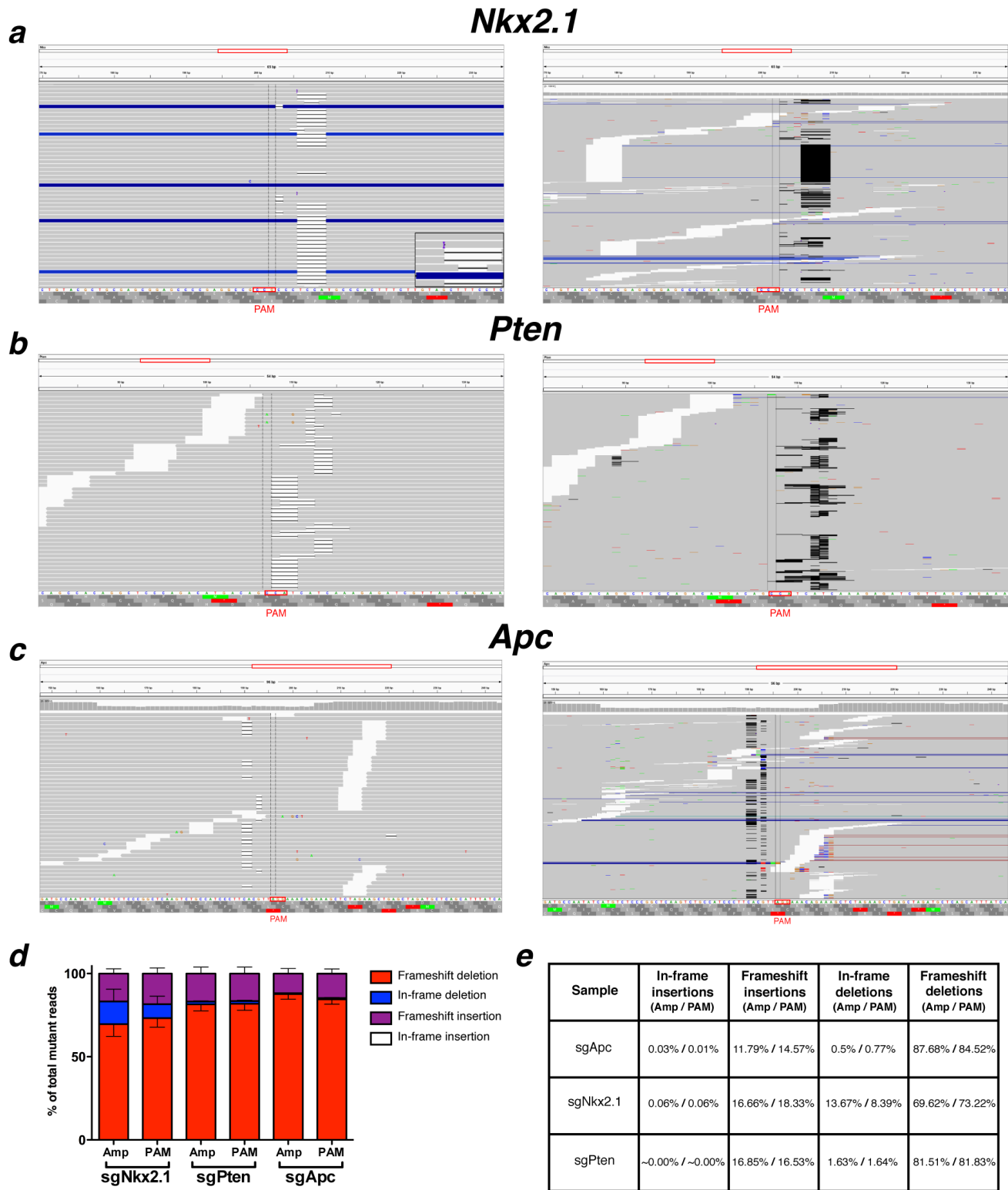
Extended Data Figure 5 | IHC-based analysis of mice infected with sgPten-pSECC. **a–c**, Negative (**a**), mixed (**b**) and positive (**c**) lung tumours of mice infected with sgPten-pSECC ($n = 9$). Positive tumour, ~100% of the tumour cells stained positive for Pten. Mixed tumour, at least ~30% of tumour

cells stained positive for Pten. Negative tumour, < 25% of the tumour cells stained positive for Pten. Dashed line in **b** demarcates the positive/negative tumour area.



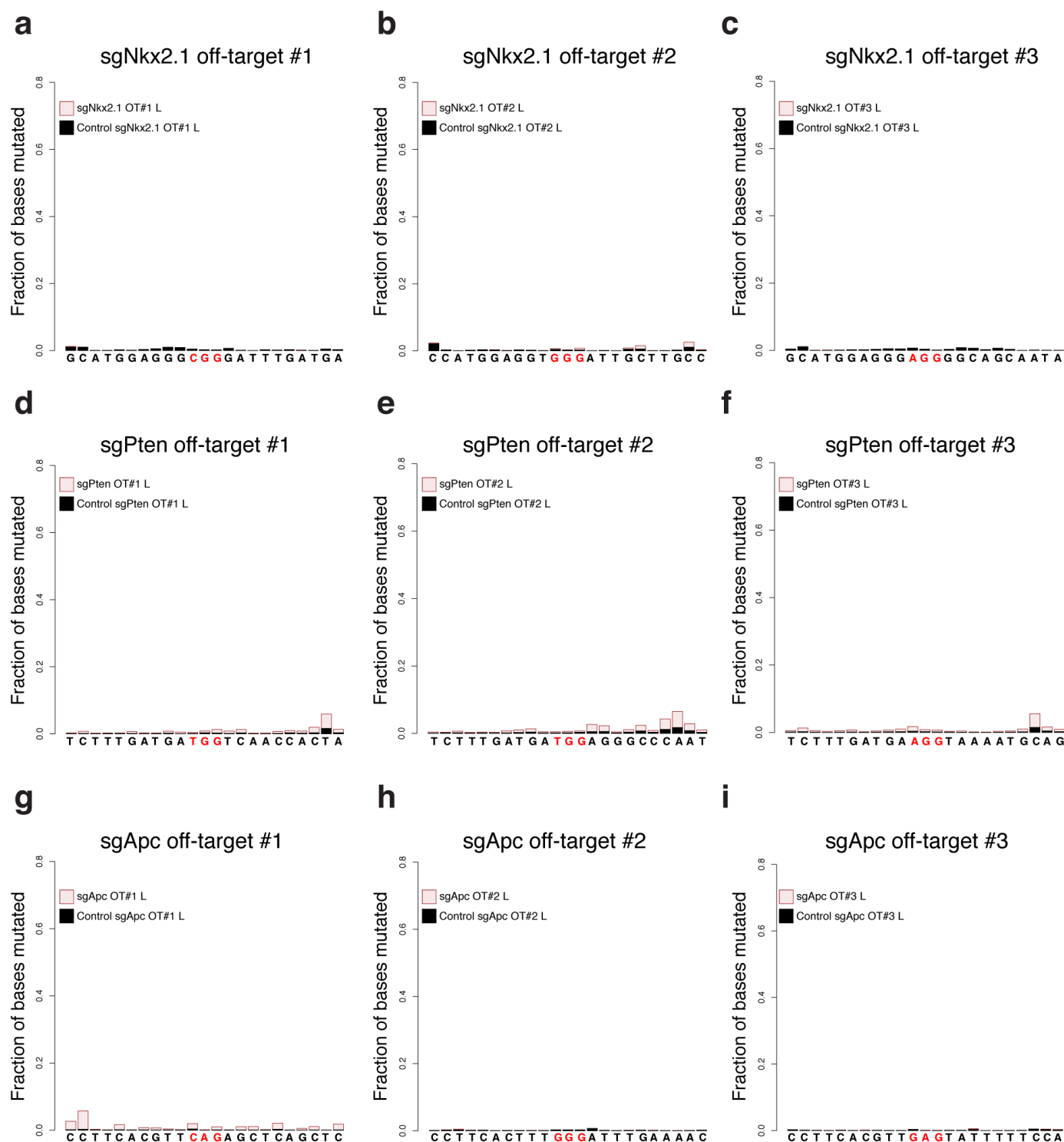
Extended Data Figure 6 | IHC-based analysis of $Kras^{LSL-G12D/+}$ and $Kras^{LSL-G12D/+}; p53^{fl/fl}$ -sgApc tumours. **a**, Representative H&E and IHC staining of serial sections from $Kras^{LSL-G12D/+}; p53^{fl/fl}$ -sgTom (control, denoted as KP-sgTom here), $Kras^{LSL-G12D/+}$ -sgApc (denoted as K-sgApc here) and $Kras^{LSL-G12D/+}; p53^{fl/fl}$ -sgApc (denoted as KP-sgApc here) lung tumours. CCSP, Clara cell secretory protein; SP-C, surfactant protein C. **b**, Contingency table demonstrating a statistically significantly higher number of β -catenin/Sox9 double-positive tumours in $Kras^{LSL-G12D/+}; p53^{fl/fl}$ -sgApc mice (29/33 tumours, 88%) vs K-sgApc mice (41/58 tumours, 71%) (one-sided chi-square test, $P < 0.05$). **c**, Percentage of all tumours that stained positive for nuclear

β -catenin that stained positive or negative for Sox9 in $Kras^{LSL-G12D/+}$ and $Kras^{LSL-G12D/+}; p53^{fl/fl}$ -sgApc mice. **d**, Contingency table demonstrating a statistically significantly higher number of tumours with Nkx2-1 low/negative areas (which are also SP-C low/negative) in sgApc-pSECC animals compared to sgTom-pSECC control animals (two-sided Fisher's exact test, $P < 0.0001$). **e**, Representative IHC staining of serial sections from an Nkx2-1 Low/Neg lung tumour obtained from a $Kras^{LSL-G12D/+}; Apc^{fl/fl}$ mouse 18 weeks after infection with Adeno-Cre. Inset shows Sox9 staining. Low/neg = tumour that had areas with clear downregulation or complete loss of Nkx2-1 or SP-C as assessed by IHC staining.



Extended Data Figure 7 | Representative examples of indels observed in lungs and tumours from mice infected with pSECC lentiviruses. a–c, Representative indels observed in the *Nkx2-1* (a), *Pten* (b) and *Apc* (c) locus from sgNkx2-1T1, sgPtenL1 and sgApcT3 samples, respectively. Left panel, details of sequence alignments around the PAM sequence. Right panel, overview of sequence alignments around the PAM sequence. Deletions and insertions are highlighted in black and purple bars, respectively. Inset in a depicts a magnification of an insertion. d, Distribution of indels (in-frame

insertions, frameshift insertions, in-frame deletions and frameshift deletions) observed in samples from mice infected with sgNkx2-1-pSECC, sgPten-pSECC and sgApc-pSECC. Amp, mutations across whole PCR amplicon; PAM, mutations across 7 base pair region upstream of the PAM sequence. e, Table summarizing percentages of indels from total mutant reads (left percentage indicates Amp (mutations across whole PCR amplicon) and right percentage indicates PAM (mutations across 7 base pair region upstream of the PAM sequence). All error bars denote s.e.m.



Extended Data Figure 8 | Off-target analysis. a–i, Analysis of off-target editing for sgNkx2-1 (a–c), sgPten (d–f) and sgApc (g–i). Briefly, potential off-target cutting at the top three predicted off-target sites (obtained from (<http://crispr.mit.edu/>); see Supplementary Table 2) for each sgRNA was assayed by Illumina MiSeq. Each plot corresponds to the fraction of bases

mutated per position in 10 bp flanks on either side of the PAM sequence (highlighted in red). Samples were obtained from entire lobes (L) from mice 10 weeks after infection with pSECC lentiviruses expressing sgNkx2-1, sgPten, sgApc or sgTom (control).

Cohesin-dependent globules and heterochromatin shape 3D genome architecture in *S. pombe*

Takeshi Mizuguchi^{1*}, Geoffrey Fudenberg^{2,3*}, Sameet Mehta¹, Jon-Matthew Belton⁴, Nitika Taneja¹, Hernan Diego Folco¹, Peter FitzGerald⁵, Job Dekker⁴, Leonid Mirny^{2,3}, Jemima Barrowman¹ & Shiv I. S. Grewal¹

Eukaryotic genomes are folded into three-dimensional structures, such as self-associating topological domains, the borders of which are enriched in cohesin and CCCTC-binding factor (CTCF) required for long-range interactions^{1–7}. How local chromatin interactions govern higher-order folding of chromatin fibres and the function of cohesin in this process remain poorly understood. Here we perform genome-wide chromatin conformation capture (Hi-C) analysis⁸ to explore the high-resolution organization of the *Schizosaccharomyces pombe* genome, which despite its small size exhibits fundamental features found in other eukaryotes⁹. Our analyses of wild-type and mutant strains reveal key elements of chromosome architecture and genome organization. On chromosome arms, small regions of chromatin locally interact to form ‘globules’. This feature requires a function of cohesin distinct from its role in sister chromatid cohesion. Cohesin is enriched at globule boundaries and its loss causes disruption of local globule structures and global chromosome territories. By contrast, heterochromatin, which loads cohesin at specific sites including pericentromeric and subtelomeric domains^{9–11}, is dispensable for globule formation but nevertheless affects genome organization. We show that heterochromatin mediates chromatin fibre compaction at centromeres and promotes prominent inter-arm interactions within centromere-proximal regions, providing structural constraints crucial for proper genome organization. Loss of heterochromatin relaxes constraints on chromosomes, causing an increase in intra- and inter-chromosomal interactions. Together, our analyses uncover fundamental genome folding principles that drive higher-order chromosome organization crucial for coordinating nuclear functions.

The 13.8-megabase (Mb) *S. pombe* genome comprises three chromosomes partitioned into euchromatin and heterochromatin domains¹². Ctr4 (known as SUV39H in mammals) and HP1 proteins assemble heterochromatin domains at pericentromeric regions, subtelomeres and the mating-type (*mat*) locus¹². Our Hi-C analysis revealed several genome organizational features (Fig. 1a).

Centromeres of all chromosomes and telomeres of chromosome 1 and 2 formed two sets of frequently interacting loci, consistent with previous work^{13,14}. Chromosome 3 ends proximal to ribosomal DNA repeats, which are compartmentalized in the nucleolus, showed no specific interactions with telomeres of chromosomes 1 and 2 (Fig. 1a). Centromeres and telomeres were refractory to interaction with chromosome arms (Fig. 1b), consistent with spatial sequestration and with similar observations in *Saccharomyces cerevisiae*¹⁵. We also found a greater frequency of inter-arm interactions than inter-chromosomal interactions, suggesting a degree of chromosome territoriality¹⁶ (Fig. 1c).

We noted a specific inter-chromosomal interaction between the right telomere of chromosome 1 (tel1R) and *mat* on chromosome 2 (Fig. 1a, b). Contact frequency was less than that between centromeres or telomeres, but was greater than the average inter-chromosomal interactions (~9-fold

enriched, Fig. 1b). Microscopy confirmed *mat*–telomere colocalization in a small proportion of cells (Fig. 1d). The dynamic inter-chromosomal *mat*–telomere interaction at the nuclear periphery^{13,17} might explain the altered intra-chromosomal pattern interactions at the *mat* locus (Extended Data Fig. 1).

We observed a notably high frequency of interactions between centromere-proximal regions, indicated by a cross-like pattern of inter-arm interactions (Fig. 1a and Extended Data Fig. 1). Previously, only direct interactions were observed between centromeres¹⁴, and a polymer modelling study did not predict this cross-like pattern for the three *S. pombe* chromosomes¹⁸. We observed similar behaviour for different chromosome arm pairs.

The polymer nature of chromatin also has an impact on genome packaging and can be studied using scaling analysis, which captures the dependency of the contact probability on genomic distance, and reflects the underlying chromatin folding status. A slow decay in contact probability at distances <100 kilobases (kb) was followed by a faster decay that falls between that of an unconstrained polymer and the fractal globule, suggesting some degree of local crumpling of the polymer⁸ (Fig. 1e and Extended Data Fig. 2). The deviation at short distances suggested additional local features of chromosome organization.

Further analyses revealed complex structures along the diagonal of the contact maps, such as at subtelomere 2R (Extended Data Fig. 1). Notably, we observed locally self-interacting domains ~50–100 kb in size, consistent with slow decay of contact probability below 100 kb (Fig. 1f). These globules were detected in all chromosomal arms. Their boundaries reflected transitions between preferential upstream and downstream interactions, and often corresponded to regions enriched for convergent genes (Fig. 1f, g). We find that globules are a prominent feature of local chromatin organization.

Cohesin affects chromatin architecture in budding yeast^{19,20} and in other eukaryotes^{5–7,21,22}, but its exact role is unclear. Cohesin enrichment at the 3' end of convergent genes (see below)²³, which correlate with globule boundaries, led us to investigate its role in globule formation. Hi-C analysis of *rad21-K1*, which contains a partial loss-of-function mutation in a cohesin subunit²⁴, revealed loss of globules and greater intermingling of chromosomes (Fig. 2a). Centromeres and telomeres were less refractory to interaction with chromosomal arms (Fig. 2b). Moreover, we observed greater intra-chromosomal inter-arm (1.6-fold increase) and inter-chromosomal contact frequencies (2.5-fold increase) compared to wild type (Fig. 2c). Contact probability decay as a function of genomic distance was quite different for *rad21-K1*. The inflection at 100 kb was absent and contact probability decayed more slowly afterwards (Extended Data Fig. 3), indicating a loss of locally compacted globules. Globule boundaries corresponded to sites of cohesin enrichment in wild type, but did not correspond to these positions in *rad21-K1* (Fig. 2d), suggesting a functional link between cohesin binding and organization of the chromatin fibre.

¹Laboratory of Biochemistry and Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. ²Graduate Program in Biophysics, Harvard University, Boston, Massachusetts 02115, USA. ³Institute for Medical Engineering and Sciences, Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ⁴Program in Systems Biology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA. ⁵Genome Analysis Unit, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA.

*These authors contributed equally to this work.

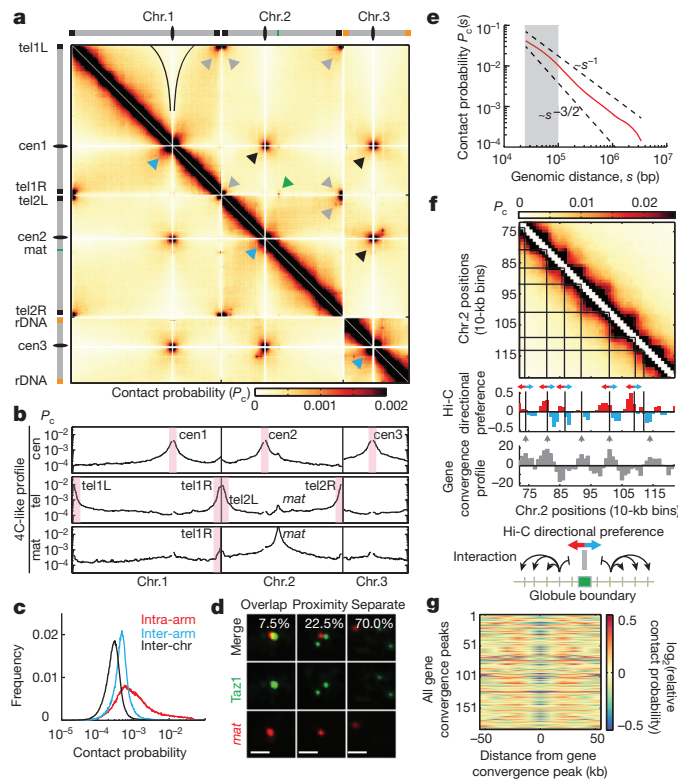


Figure 1 | High-resolution contact probability map reveals the higher-order organization of the *S. pombe* genome. **a**, Genome wide Hi-C heatmap at 10-kb resolution. Grey and black arrowheads: telomere and centromere clustering, respectively; blue arrowheads: centromere proximal arm-arm interactions; green arrowhead: *mat*–tel1R interaction. Black outlines indicate centromere avoidance of chromosome (chr.) arms. P_c , contact probability. **b**, 4C-like inter-chromosomal interaction profiles showing the average contact probabilities of centromeres and telomeres. Inter- and intra-chromosomal interaction profiles are shown for *mat*. **c**, Distribution of intra-arm, inter-arm and inter-chromosomal contact frequencies. Inter-arm interactions were 1.7-fold higher than inter-chromosomal interactions for regions >200 kb from centromeres and telomeres. **d**, Visualization of telomeric Taz1 and *mat* by immunofluorescence and FISH ($n = 400$ nuclei). Scale bars, 1 μ m. **e**, Decay of intra-arm contact probability as a function of genomic distance, s ($P_c(s)$). $P_c(s)$ decreases more slowly at short distances (grey shaded area). Dashed lines represent the slopes for polymers in a melt ($-3/2$) and fractal globules (-1). **f**, Hi-C heatmap for a sub-chromosomal region showing globules. The Hi-C directional preference profile is shown, with upstream (red) and downstream (blue) preferences. The gene convergence profile is shown underneath. Grey arrows: peaks of convergent gene enrichment; black lines: globule boundaries. **g**, Relative contact probability averaged over 20–50 kb for all gene convergence peaks. Decreased relative contact probability at peaks indicates that regions on either side of the peak are insulated from each other.

We next examined the relationship between globules and cohesin profiles binned to 10-kb resolution as for Hi-C analysis. First, we measured average insulation around cohesin peaks by calculating the relative contact probability at a given genomic distance (Extended Data Fig. 4a, b). Contact frequency between regions separated by cohesin peaks was depleted in wild type, and this depletion was lost in *rad21-K1*, suggesting a cohesin-dependent interaction barrier with an effective range of ~50–100 kb (Fig. 2e). Second, insulation analyses at each cohesin peak showed that cohesin-mediated insulation is a general feature of wild type but not *rad21-K1* (Extended Data Fig. 4c). Third, we determined the mean number of cohesin peaks as a function of distance to the nearest boundary between preferential upstream/downstream interactions. Cohesin peaks were enriched at boundaries specifically in wild type (Extended Data Fig. 4d). Thus, cohesin maintains globule boundary positions throughout the genome. Finally, a genome-wide correlation

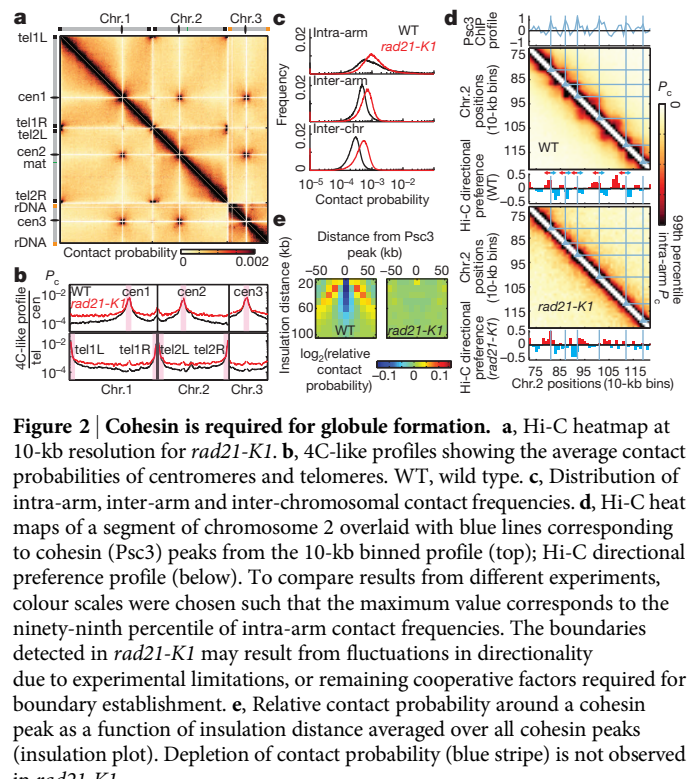


Figure 2 | Cohesin is required for globule formation. **a**, Hi-C heatmap at 10-kb resolution for *rad21-K1*. **b**, 4C-like profiles showing the average contact probabilities of centromeres and telomeres. WT, wild type. **c**, Distribution of intra-arm, inter-arm and inter-chromosomal contact frequencies. **d**, Hi-C heat maps of a segment of chromosome 2 overlaid with blue lines corresponding to cohesin (Psc3) peaks from the 10-kb binned profile (top); Hi-C directional preference profile (below). To compare results from different experiments, colour scales were chosen such that the maximum value corresponds to the ninety-ninth percentile of intra-arm contact frequencies. The boundaries detected in *rad21-K1* may result from fluctuations in directionality due to experimental limitations, or remaining cooperative factors required for boundary establishment. **e**, Relative contact probability around a cohesin peak as a function of insulation distance averaged over all cohesin peaks (insulation plot). Depletion of contact probability (blue stripe) is not observed in *rad21-K1*.

between the profile of cohesin enrichment and the depletion of inter-actions between globules observed in wild type for up to 100 kb was absent in *rad21-K1*, suggesting that both the position and amount of cohesin contribute to boundary function (Extended Data Fig. 4e). Additional factor(s) may also determine globule boundaries.

We then considered the functional importance of cohesin-dependent globules. We found local duplications at loci with high sequence similarity, such as retrotransposons, long terminal repeats and pericentromeric repeats (Extended Data Fig. 5a). Thus, constraints imposed by cohesin may prevent ectopic recombination between repeats. Because defective cohesin also impairs transcription termination at select convergent genes²⁵, we wondered whether cohesin-mediated genome organization acts broadly to restrict inappropriate RNAPII activity. Expression profiling indeed revealed widespread read-through transcripts in *rad21-K1* (Extended Data Fig. 5b).

Asynchronous *S. pombe* cultures contain most cells in G2, in which cohesin is required for sister chromatid cohesion. To determine whether globules are also present in G1, we performed Hi-C and cohesin mapping using the cell cycle mutant *cdc10-v50* arrested in G1 (Fig. 3a). Consistent with previous work²⁶, we detected cohesin in G1 cells (Extended Data Fig. 6a), with prominent enrichment at convergent genes (Fig. 3b, c). The Hi-C contact map was similar to asynchronous cells (Extended Data Fig. 6b, c). Importantly, we detected globules in G1 cells, consistent with slow decay of contact probability at short distances (Fig. 3d and Extended Data Fig. 6d). Moreover, cohesin is required for maintenance of globules and inflections in scaling in G1-arrested cells (Extended Data Fig. 6e, f). These results suggest that key features of genome architecture are preserved in G1.

We further analysed G1 globule boundaries. Average insulation around cohesin peaks at globule boundaries indicated that they function as interaction barriers only slightly less efficiently than in asynchronous cells (Fig. 3e). Interaction barrier function was eliminated in G1-arrested *rad21-K1* (Extended Data Fig. 6g). Cohesin enrichment inversely correlated with relative contact probability genome-wide, and globule boundary positions overlapped with wild type (Extended Data Fig. 7a–c). Concentrated cohesin might create barriers for local chromatin compaction

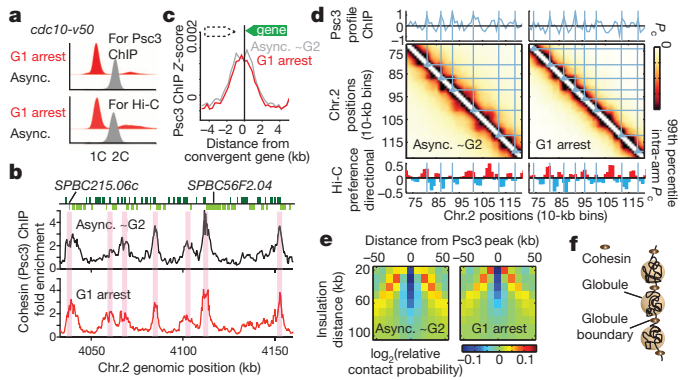


Figure 3 | Globules are a feature of chromosome architecture in G1 cells. **a**, FACS of cells used for cohesin ChIP and Hi-C. **b**, Cohesin (Psc3) ChIP enrichment at convergent genes within a region of chromosome 2. **c**, Psc3 enrichment at convergent genes in wild-type and G1 cells. All pairs of convergent genes were aligned at the 3' end of the second gene (green box). **d**, Hi-C heatmaps of a segment of chromosome 2 overlaid with lines corresponding to cohesin peaks from the 10-kb binned profile. Plotted below is the directional preference profile. **e**, Insulation plot around cohesin peak in G1-arrested cells. **f**, Hypothetical model showing cohesin bound between locally compacted globule domains. Cohesin confines interactions within individual globule domains and prevents interactions across the boundaries.

factors, or the cohesin ring²⁷ might constrict borders to create globules (Fig. 3f). Importantly, we find globules are a feature of both G1 and G2 genome architecture.

Heterochromatin facilitates cohesin binding at specific loci^{10,11,23,25} and may affect genome organization²⁸. Hi-C analysis of a strain lacking the sole H3K9 methyltransferase Clr4, which is required for heterochromatin assembly¹², revealed widespread changes (Fig. 4a). Heterochromatic regions (centromeres, telomeres and *mat*) were less refractory to genome-wide interactions (Fig. 4b). We observed strong interactions

between *mat* and all telomeres in *clr4Δ* (Fig. 4b), confirmed by microscopy (54% overlap/proximity). The increased intra-chromosomal inter-arm interactions and inter-chromosomal interactions in *clr4Δ* (Fig. 4c) are consistent with reduced chromosome territoriality.

Defective cohesin loading could cause global changes in *clr4Δ*. We observed a major reduction in cohesin at pericentromeric and subtelomeric domains (Extended Data Fig. 8a, b), but not at chromosomal arms where cohesin peaks correlated with convergent gene enrichment (Extended Data Fig. 8c). Consistent with this, globules were not affected in *clr4Δ*, and cohesin enrichment was coincident with globule boundaries (Fig. 4d, e and Extended Data Fig. 7a). Cohesin enrichment and relative contact probability were inversely correlated in *clr4Δ*, and globule boundaries overlapped with wild type (Extended Data Fig. 7b, c). Thus, the loss of heterochromatin does not affect globules along chromosomal arms.

Notably, the cross-like pattern of centromere-proximal interactions was less evident in *clr4Δ* (Fig. 4a, b, f). These contacts were diminished, as were telomere–telomere contacts (Extended Data Fig. 9). We used a modified scaling method to further examine contact probabilities between centromere-proximal arm regions of the same and different chromosomes. We determined contact probability scaling between arm pairs as a function of genomic separation, defined for two loci as the sum of their respective distances from the centromere (Extended Data Fig. 10a). In wild type, inter-arm intra-chromosomal scaling, between two arms of the same chromosome, was very similar to the intra-arm scaling (Extended Data Fig. 10b). Inter-arm inter-chromosomal scaling, between arms of different chromosomes, was also similar, although shifted slightly lower, indicating a consistently lower contact frequency than for intra-chromosomal inter-arm interactions (Extended Data Fig. 10b). In *clr4Δ*, inter-arm scalings were shifted lower for arms of the same and different chromosomes, indicating lower contact frequency and greater distance between arms extending from centromeres (Extended Data Fig. 10c). This decrease is not due solely to the loss of pericentromeric cohesin, as inter-arm contact between chromosomes in *rad21-K1* was

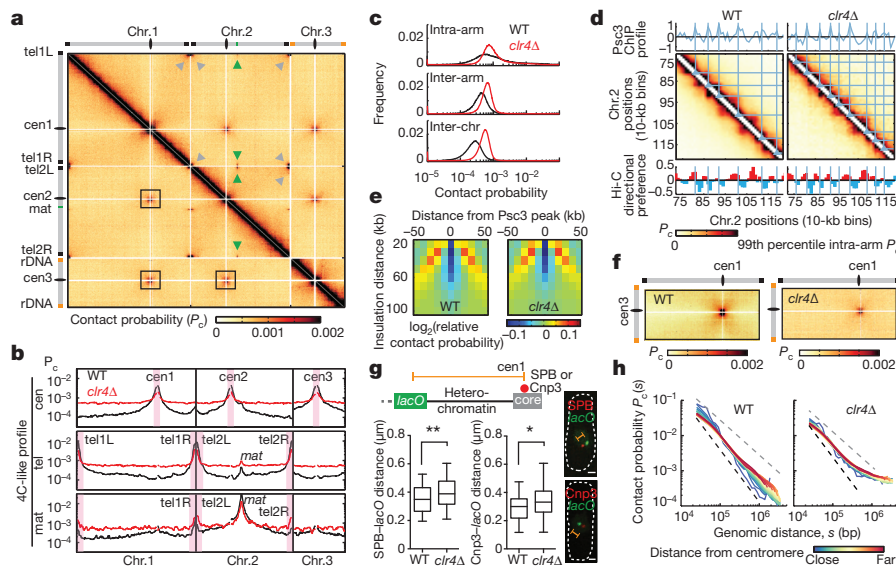


Figure 4 | Loss of heterochromatin affects genome organization. **a**, Hi-C heatmap for *clr4Δ*. **b**, 4C-like interaction profiles showing the average contact probabilities of centromeres and telomeres. Inter- and intra-chromosomal interaction profiles are shown for *mat*. **c**, Distribution of indicated contact probabilities. **d**, Hi-C heatmaps of a segment of chromosome 2 overlaid with lines corresponding to cohesin peaks. Directional preference profiles are plotted below. **e**, Insulation plot around cohesin peak in *clr4Δ*. **f**, The cross-like pattern of interactions between centromere-proximal regions of different chromosomes. **g**, Box and whisker plot showing the distribution of spatial distances between *lacO* inserted outside a pericentromeric heterochromatin

domain and the centromere core marked by either SPB (left) or Cnp3 (right). * $P \leq 0.05$, ** $P \leq 0.01$ ($n = 90$, two-sided Mann–Whitney U test).

Scale bars, 1 μm . **h**, Intra-arm $P_c(s)$ for regions coloured according to their distance from the centromere. In wild type, $P_c(s)$ for regions near the centromere decreases more rapidly than for more distal regions. In *clr4Δ*, decay of $P_c(s)$ for regions near the centromere is similar to that of more distal regions, indicating that the organization and conformations of pericentromeric chromatin are more similar to the organization of other chromosomal regions.

similar to wild type (Extended Data Fig. 10c), suggesting that heterochromatin itself organizes pericentromeric regions.

The changes observed in *clr4Δ* are consistent with an increase in the contour length of pericentromeric regions, resulting from loss of compacted heterochromatin. To explore this, we studied the effect of *clr4Δ* on chromatin at a pericentromeric domain. Consistent with scaling analysis of *clr4Δ*, distance distributions increased between the centromere core and a *lacO* array inserted outside the heterochromatin domain (Fig. 4g). These results indicate that chromatin fibre compaction by heterochromatin imposes additional structural constraints.

Our results suggested co-linear extension of centromere proximal regions, with less constrained distal regions. To determine whether loss of heterochromatin relaxes these constraints we performed scaling analysis of individual chromosome segments at increasing distance from the centromere. In wild type, contact probability decayed more rapidly for chromosome segments near centromeres than for those farther away, consistent with tight clustering of centromeres and a volume exclusion effect (Fig. 4h). By contrast, contact probability decay was similar for all chromosome segments in *clr4Δ*, regardless of distance from the centromere (Fig. 4h). These results indicate that chromatin compaction at pericentromeric regions promotes spatial restriction of the genome.

Our results reveal two new aspects of chromatin organization: 50–100-kb globules and strong heterochromatin-mediated interactions between centromere-proximal regions (Extended Data Fig. 10d). Globules require cohesin and are a basic element of chromosome arm architecture, distinct from cohesin-dependent long-range loop interactions between gene regulatory elements in higher eukaryotes^{5,21,22}. Globules may be integral components of larger domains in other species, and could explain changes within topological domains after cohesin depletion^{6,7}. Globules comprising crumpled chromatin may facilitate functional genome annotation and promote transcriptional fidelity. Heterochromatin imposes an additional constraint, perhaps partly mediated by cohesin at centromeres and telomeres, and compacts large domains at opposing ends of the nucleus that may help reinforce Rabl organization. Heterochromatin-mediated condensation and globule assembly probably fulfil complementary roles to constrain chromosomal arms and promote territoriality. These results uncover distinct aspects of genome architecture, and lay the groundwork for future investigation of its impact on various chromosomal processes.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 September 2013; accepted 5 September 2014.

Published online 12 October 2014.

- Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
- Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
- Sexton, T. *et al.* Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).
- Hou, C., Li, L., Qin, Z. S. & Corces, V. G. Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol. Cell* **48**, 471–484 (2012).
- Phillips-Cremins, J. E. *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281–1295 (2013).
- Sofueva, S. *et al.* Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J.* **32**, 3119–3129 (2013).

- Zuin, J. *et al.* Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl Acad. Sci. USA* **111**, 996–1001 (2014).
- Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
- Grewal, S. I. & Jia, S. Heterochromatin revisited. *Nature Rev. Genet.* **8**, 35–46 (2007).
- Bernard, P. *et al.* Requirement of heterochromatin for cohesion at centromeres. *Science* **294**, 2539–2542 (2001).
- Nonaka, N. *et al.* Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. *Nature Cell Biol.* **4**, 89–93 (2002).
- Cam, H. P. *et al.* Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nature Genet.* **37**, 809–819 (2005).
- Funabiki, H., Hagan, I., Uzawa, S. & Yanagida, M. Cell cycle-dependent specific positioning and clustering of centromeres and telomeres in fission yeast. *J. Cell Biol.* **121**, 961–976 (1993).
- Tanizawa, H. *et al.* Mapping of long-range associations throughout the fission yeast genome reveals global genome organization linked to transcriptional regulation. *Nucleic Acids Res.* **38**, 8164–8177 (2010).
- Duan, Z. *et al.* A three-dimensional model of the yeast genome. *Nature* **465**, 363–367 (2010).
- Molnar, M. & Kleckner, N. Examination of interchromosomal interactions in vegetatively growing diploid *Schizosaccharomyces pombe* cells by Cre/loxP site-specific recombination. *Genetics* **178**, 99–112 (2008).
- Alfredsson-Timmings, J., Henningson, F. & Björling, P. The *Clr4* methyltransferase determines the subnuclear localization of the mating-type region in fission yeast. *J. Cell Sci.* **120**, 1935–1943 (2007).
- Tjong, H., Gong, K., Chen, L. & Alber, F. Physical tethering and volume exclusion determine higher-order genome organization in budding yeast. *Genome Res.* **22**, 1295–1305 (2012).
- Gard, S. *et al.* Cohesinopathy mutations disrupt the subnuclear organization of chromatin. *J. Cell Biol.* **187**, 455–462 (2009).
- Guacci, V., Koshland, D. & Strunnikov, A. A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in *S. cerevisiae*. *Cell* **91**, 47–57 (1997).
- Dorsett, D. & Merkenschlager, M. Cohesin at active genes: a unifying theme for cohesin and gene expression from model organisms to humans. *Curr. Opin. Cell Biol.* **25**, 327–333 (2013).
- Seitan, V. C. *et al.* Cohesin-based chromatin interactions enable regulated gene expression within pre-existing architectural compartments. *Genome Res.* **23**, 2066–2077 (2013).
- Schmidt, C. K., Brookes, N. & Uhlmann, F. Conserved features of cohesin binding along fission yeast chromosomes. *Genome Biol.* **10**, R52 (2009).
- Tomonaga, T. *et al.* Characterization of fission yeast cohesin: essential anaphase proteolysis of Rad21 phosphorylated in the S phase. *Genes Dev.* **14**, 2757–2770 (2000).
- Gullerova, M. & Proudfoot, N. J. Cohesin complex promotes transcriptional termination between convergent genes in *S. pombe*. *Cell* **132**, 983–995 (2008).
- Bernard, P. *et al.* Cell-cycle regulation of cohesin stability along fission yeast chromosomes. *EMBO J.* **27**, 111–121 (2008).
- Haering, C. H., Farcas, A. M., Arumugam, P., Metson, J. & Nasmyth, K. The cohesin ring concatenates sister DNA molecules. *Nature* **454**, 297–301 (2008).
- Renauld, H. & Gasser, S. M. Heterochromatin: a meiotic matchmaker? *Trends Cell Biol.* **7**, 201–205 (1997).

Acknowledgements We thank Y. Watanabe and M. Yanagida for strains, M. Zofall for contributions, and A. Kelly for comments. This study used the Helix Systems and the Biowulf Linux cluster at the National Institutes of Health. This work was supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute and by a grant from NHGRI (HG003143) to J.D. The work of G.F. and L.M. is supported by NCI Physical Sciences-Oncology Center at MIT (U54CA143874).

Author Contributions T.M. and S.I.S.G. designed the experiments. T.M. did most experiments including Hi-C with guidance from J.-M.B. and J.D.; G.F., S.M. and L.M. performed the Hi-C computational analyses; P.F. contributed to data analyses; N.T. and H.D.F. provided imaging analyses. All authors contributed to data interpretation. T.M., J.B. and S.I.S.G. wrote the manuscript with input from all authors.

Author Information Microarray and sequencing data are available at the NCBI Gene Expression Omnibus (GEO) repository under the accession number GSE56849. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.I.S.G. (grewals@mail.nih.gov).

METHODS

Hi-C. For Hi-C experiments, wild-type and *clr4Δ* strains were cultured in rich medium at 33 °C. The *cdc10-v50*, *rad21-K1* and *cdc10-v50 rad21-K1* mutant cells were initially cultured at 26 °C and then shifted to 35 °C for 4 h (*cdc10-v50* and *cdc10-v50 rad21-K1*) or to 33 °C for only 2 h (*rad21-K1*) to ensure that the cell cycle distribution of mutant cultures was similar to wild type. A detailed Hi-C protocol was described previously²⁹. In brief, cells ($A_{600\text{ nm}} \sim 0.5$) were fixed in 3% formaldehyde (Sigma) for 20 min at 26 °C, and quenched with glycine for 5 min at 26 °C. Cells were poured into liquid nitrogen using NEBuffer2 (New England Biolabs), disrupted by nitrogen grinding. Cell lysate was treated with 0.1% SDS for 10 min at 65 °C, and then quenched with 1% Triton X-100. Cell lysate was digested overnight with HindIII at 37 °C. The 5' overhang from HindIII digestion was filled in using the Klenow fragment in the presence of biotin-14-dCTP, dATP, dGTP and dTTP at 37 °C for 45 min. The reaction was terminated with 1.5% SDS. The DNA fragments were ligated by T4 DNA ligase in diluted conditions that favour the ligation between cross-linked DNA fragments at 16 °C for 8 h (Hi-C DNA). The Hi-C DNA was reverse cross-linked at 65 °C overnight in the presence of proteinase K and purified by phenol/chloroform extraction. Purified Hi-C DNA was treated with 1 mg ml⁻¹ RNase A for 30 min at 37 °C. Biotinylated but not ligated DNA fragments were removed by T4 DNA polymerase and reactions were purified by phenol/chloroform extraction. Hi-C DNA was then sheared using the Covaris S2 instrument (Covaris) in the size range of <500 bp. The sheared Hi-C DNA was subjected to end-repair and 3' end adenylation. Hi-C DNA between 150 and 300 bp was selected with AMPure XP (BeckmanCoulter) as described²⁹. The biotin-labelled Hi-C DNA was selectively captured by Dynabeads Myone Streptavidin C1 (Invitrogen) and used for Illumina PE adaptor ligation. Streptavidin beads containing bound Hi-C DNA were used for the template for library amplification by PE-PCR primers (Illumina). Hi-C libraries were sequenced using the Illumina HiSeq platform. Analysis of biological replicates for wild type and mutants in our laboratory give similar contact probability results. The correlation values between two wild type samples yielded Pearson's coefficient $r = 0.981$, $P < 2 \times 10^{-16}$.

Hi-C data analyses overview. Hi-C data were mapped, and reads were filtered as described previously³⁰. Corrected contact probability matrices at 10-kb resolution were obtained using iterative correction³⁰. Both steps were performed using the hiclib library for python, publicly available at <https://bitbucket.org/mirnylab/hiclib>.

Mapping and fragment-level filtering. Paired-end sequencing reads were mapped independently using Bowtie 2.1.0 to the *S. pombe* reference genome (ASM294v2) for each Hi-C library. Mapping with iteratively increasing truncation length was used to maximize yield of valid Hi-C interactions, using the flags '-score-min L,-0.6,-0.2', as described previously³⁰. Only read pairs where both reads uniquely aligned to the genome were considered for subsequent steps. Read pairs corresponding to repeat instances of the same DNA molecule were removed. Next, on the basis of their HindIII restriction fragment assignments and orientations, read pairs were classified as valid Hi-C products, non-ligation, or self-ligation products³⁰. The following fragment-level filters were then applied, as described³⁰, which remove read pairs: with one end adjacent to the restriction site (possible unligated molecules), from restriction fragments with very high or low counts, from very large or small restriction fragments, and separated by very few restriction fragments (as these may be strongly influenced by inefficiencies in restriction). Filters used hiclib default values, except for the last filter, which used a more stringent 4 instead of 2 fragments. For downstream analyses, only valid Hi-C read-pairs were considered. Furthermore, read pairs from biological replicates were pooled after applying fragment-level filters. The number of valid read pairs used were as follows: 61,873,904 for wild type; 12,521,720 for *rad21-K1*; 16,821,386 *cdc10-v50* (G1 arrested cells); and 18,549,406 for *clr4Δ*.

Corrected Hi-C contact maps. To create contact maps, the *S. pombe* genome was divided into non-overlapping 10-kb bins. We then assigned valid Hi-C products to the bins based on the midpoint of the associated restriction fragment, as previously^{8,30}. As previously described³⁰, we used bin-level filters to focus our analyses on regions of the genome that could be most reliably assessed with Hi-C, removing: the lowest 1% of bins by coverage (in addition to bins with zero counts), the diagonal and neighbouring diagonal (that is, bin pairs separated by <20 kb), stand-alone bins (that is, bins in which both neighbouring bins did not pass filters). We then removed potential biases in raw Hi-C contact maps, which may include the uneven distribution of restriction enzyme sites, differences in GC content, and differing mappability of different bins. This was achieved by normalizing coverage using an iterative procedure³⁰. Regions of the heatmap in which a single bin had been filtered out were then interpolated using neighbouring bins: within a chromosome, position (i, j) was interpolated with the average value at positions ($i+1, j+1$) and ($i-1, j-1$) to preserve the decrease of contact probability $P_c(s)$ with distance s ; at the edges of chromosomes, if the offset in i or j changed chromosomal assignment, the average of ($i+1, j$) and ($i-1, j$) or ($i, j+1$) and ($i, j-1$) was used instead; the latter was also used between chromosomes, except at the intersection of two interpolated

bins, where the average of ($i+1, j+1$), ($i-1, j-1$), ($i+1, j-1$) and ($i-1, j+1$) was used. The resulting matrices were then normalized so that each row and column sum to 1. All reported Hi-C results use these normalized corrected matrices.

4C-like profiles. To obtain 4C-like inter-chromosomal interaction profiles for centromeres and telomeres (for example, Fig. 1b) indices of telomeric and centromeric bins were extended to include the nearest 5 non-filtered out bins on their respective chromosomes. The profiles for each of these sets of regions were then averaged together over all non-filtered and non-intra-chromosomal and smoothed with a sliding window of 5 bins. For the *mat* locus, intra-chromosomal interactions were plotted as well, making it an exact analogue of a 4C profile obtained from corrected Hi-C data.

$P_c(s)$ calculation. Polymers characteristically display a decrease in contact probability, $P_c(s)$, as a function of genomic distance, s . The rate of decay, or scaling, of $P_c(s)$ is often interpreted as informative of an underlying polymer state. In particular, $P_c(s) \approx s^{-1}$ has been interpreted as indicative of a non-equilibrated crumpled, or fractal, globule state, which stands in contrast with $P_c(s) \approx s^{-3/2}$ for a polymer melt³¹. Here, we observe that mutants often display markedly different $P_c(s)$ values than wild-type *S. pombe*, which decreases at a rate inbetween s^{-1} and $s^{-3/2}$ after 100 kb. Intra-arm $P_c(s)$ values (for see, example, Fig. 1e) was calculated from binned corrected contact maps as described previously³⁰, in which intra-arm regions were defined as all pairs of bins on the same arm of the same chromosome. Centromere coordinates from the reference *S. pombe* genome (ASM294v2) were used to define chromosomal arms. First, we consider 40 logarithmically spaced bins from 20 kb to the maximum arm length; bin positions were rounded down to the nearest integer and repeated bins locations were discarded. For each logarithmic bin, we then calculate the mean value of Hi-C contact map in this range of genomic distances, excluding regions of the contact map that were filtered out. To determine intra-arm $P_c(s)$ as function of the distance of a region to the centromere (Fig. 4h), we assigned regions to 20 logarithmically spaced bins as a function of distance to the centromere. The intra-arm analysis was then performed separately for each of these sets of regions. Inter-chromosomal $P_c(s)$ as a function of combined distance to the centromere was calculated similarly to intra-arm $P_c(s)$, with two important differences. First, instead of s representing genomic separation between two loci, $s = d_1 + d_2$, in which d_1 and d_2 are the respective distances to the centromere of the first and second arm for each inter-chromosomal arm pair. Second, this $P_c(s)$ calculation was restricted to loci at similar distances from the centromere, $|d_1 - d_2| < 50$ kb.

Histograms of contact probability. Histograms of contact probability for different classes of regions (for example, Fig. 1c) were calculated from corrected contact maps, excluding filtered-out bin pairs. Inter-chromosomal regions were defined as all pairs of bins on different chromosomes. Inter-arm regions were defined as all pairs of bins on different arms of the same chromosome. As for calculating $P_c(s)$, intra-arm regions were defined as all pairs of bins on the same arm of the same chromosome.

Hi-C directional preferences and globule boundaries. Hi-C directional preference scores were calculated from corrected contact maps as the log₂ ratio of upstream to downstream contact probabilities for each region at distances below 100 kb:

$$D_i = \log_2 \left(\sum_{j=-10}^j C_{i,i+j} / \sum_{j=0}^{j=10} C_{i,i+j} \right)$$

in which C is the corrected contact map. Globule boundaries occur where the directional preferences strongly change from regions of upstream preferences to regions of downstream preferences. Boundary strength was calculated as the sum of upstream preferences in the region before, minus the sum of downstream preferences in the region after, a boundary. For comparisons of boundary position between data sets, and comparisons with positions of peaks of local cohesin enrichment, the 100 strongest boundaries in each data set were used.

Gene convergence profile. We constructed a gene convergence profile, $L_{10\text{ kb}}(i)$, at 10-kb resolution in three steps: (1) the gene orientation profile, $G_{1\text{ kb}}(i)$, was calculated at 1-kb resolution, in which $G_{1\text{ kb}}(i) = 1, -1$ or 0 for downstream, upstream or no orientation, respectively, and genes were assigned to bins by their midpoint. The value in each 1-kb bin was set to 1 if the number of downstream genes exceeded the number of upstream, -1 if the number of upstream genes exceeded the number of downstream and 0 otherwise. (2) The gene orientation profile $G_{1\text{ kb}}(i)$ was then used to calculate the convergence profile $L_{1\text{ kb}}(i)$ at the same 1-kb resolution. For each position i the convergence is calculated as the weighted sum of positive gene orientation bins 50-kb upstream and negative bins 50-kb downstream:

$$L_{1\text{ kb}}(k) = \sum_{j=-50}^{j=50} W(j) G_{1\text{ kb}}(k+j)$$

in which $W(j) = \text{sign}(j) \frac{(|j| - 50)}{50}$. The triangular shape of kernel $W(j)$ that changes sign at 0 allows smooth weighting of the directionality of genes as a function of distance. The convergence profile $L_{1\text{ kb}}$ has positive values for regions of convergent gene orientation, that is, more upstream genes point upstream and more downstream genes point downstream. Conversely, negative values of $L_{1\text{ kb}}$ indicate regions of divergent gene orientation. (3) The 10-kb profile $L_{10\text{ kb}}(i)$ is calculated by averaging the 1-kb scores in each non-overlapping 10-kb bin. This allows comparison with Hi-C contact maps binned at 10 kb (Fig. 1f). Finally, peaks of convergent gene orientation at 10 kb were defined as bins in the top seventy-fifth percentile of $L_{10\text{ kb}}$ separated by at least three bins, as for peaks of local cohesin enrichment, and represent when gene orientation shifts from mostly downstream to mostly upstream at the 100-kb scale.

Binned cohesin analysis and peak detection. To compare Hi-C contact maps binned at 10 kb with cohesin (Psc3) binding, we constructed a cohesin profile and obtained regions of local cohesin enrichment at this scale for the wild type and for each mutant. First, log ratio ChIP-Chip values were averaged over 10-kb non-overlapping bins to obtain the 10-kb profile. Next, peaks of local cohesin enrichment were determined as local maxima in the 10-kb binned profile. Peaks were additionally required to have a minimum spacing of 3 bins and to be in the top seventy-fifth percentile.

Relative contact probability/insulation calculation. The degree to which a locus can decrease contact frequency between, or insulate, regions separated by that locus can be directly quantified from corrected Hi-C contact maps. To quantify the relative frequency of contacts occurring over a bin j at a distance s , we calculate:

$$R_j(s) = \log_2 \left(\sum_{k=j-s/2}^{k=j+s/2} C_{k,k+s} / M \right)$$

in which $M = \text{mean}_s \left(\sum_{k=j-s/2}^{k=j+s/2} C_{k,k+s} \right)$. $R_j(s)$ provides a natural way to determine whether certain regions are associated with insulation in a Hi-C contact map; negative values of $R_j(s)$ indicate fewer contacts occurring over given bin, that is, insulation, at a given distance s .

Insulation versus peaks. To plot the local relative contact probability around all convergent gene peaks (Fig. 1g) or Psc3 peaks (Extended Data Fig. 4c), $R_j(s)$ was averaged over $s = 20\text{--}50$ kb, and $R_j(s) - \text{mean}_{j-5}^{j+5} R_j(s)$ was plotted for ± 5 bins (50 kb) from each peak location j . To plot the average local relative contact probability profile surrounding Psc3 peaks as a function of genomic distance, $R_j(s) - \text{mean}_{j-5}^{j+5} R_j(s)$ was averaged over all peak locations, j , and $R(s)$ was plotted for ± 5 bins (50 kb) offset (for example, Fig. 2e).

Cohesin peaks and distance to the nearest boundary. To calculate the mean number of cohesin/Psc3 peaks as a function of distance to the nearest boundary at 10 kb (Extended Data Fig. 4d), we used wild-type Psc3 peaks as determined above from the 10-kb binned profile, and the top 100 strongest boundaries for each data set. Error bars showing fifth and ninety-fifth percentiles were obtained by comparison with Psc3 peaks determined from 1,000 permuted Psc3 profiles. The Spearman correlation between cohesin and relative contact frequency at 10 kb as a function of distance was calculated genome-wide as the correlation between $R(s)$ and the binned Psc3 profile at each distance s (Extended Data Figs 4e and 7b). Negative values of this correlation indicate that with increased cohesin binding relatively fewer contacts are made over a locus at a given distance.

Box plots of inter-chromosomal contact probability. Box plots of inter-chromosomal interactions for telomere–telomere interactions show the 10 most telomere-proximal bin-pairs for non-filtered regions of the heatmap. For centromere–centromere interactions the values for the 40 most centromere-proximal bin-pairs for non-filtered regions of the heatmap are used, as centromere–centromere interactions occur where four distinct arm pairs meet. As contact probabilities span a large range, $\log_{10}(\text{contact probability})$ is shown (Extended Data Fig. 9).

Average cohesin profile at convergent genes. For each sample the log ratio data were averaged over a 50-bp sliding window. All pairs of convergent genes were aligned at the 3' end of the second gene in the pair and the mean (geometric mean) of all genes for positions $\pm 5,000$ bases from the align-point was plotted. Gene boundaries (start and end positions) were as previously reported³². To calculate z-score profile for cohesin binding at convergent genes, initial cohesin binding data was mapped to genomic coordinates and represented as per-probe log-ratios. For comparison between different experiments the log-ratio value of each probe was converted to a z-score. Using the coordinates of the convergent genes we took a window of ± 5 kb around one 3' end of the convergent gene pair. This window was

then divided into 300-bp bins. Thus, for each 10-kb window we have 30 bins, and each window is now centred around the 3' end of the right-hand gene of each convergent gene pair. For each base lying within the 300-bp bin the z-scores are averaged and assigned to the midpoint of the 300-bp bin. Finally, the score of each bin was summed for 1,463 genes across the 10-kb windows and plotted as the mean score per gene (Fig. 3c).

Immunofluorescence and FISH. Immunofluorescence/fluorescent *in situ* hybridization (FISH) was carried out as previously described³³. Rabbit anti-GFP (ab290, Abcam) and anti-TAT1 antibodies (gift from K. Gull) were used for detecting the Psc3–GFP and tubulin, respectively. Rabbit anti-GFP (ab290, Abcam) was used to detect Taz1–GFP. Cells were subsequently incubated with Alexa Fluor 488 anti-rabbit IgG (Molecular Probes, Invitrogen). The FISH probe was designed by HindIII digestion of plasmids containing various *mat1*, *mat2* and *mat3* loci. The purified products were digoxigenin (DIG)-labelled using a Nick Translation kit (Roche Applied Science). Cells were hybridized overnight with DIG-labelled probe and signals were detected using Fab fragments from polyclonal anti-digoxigenin antibodies conjugated to rhodamine (Roche Applied Science). Samples were analysed using a Delta Vision Elite fluorescence microscope with oil immersion objective lens of $\times 100$ magnification, numerical aperture (NA) 1.4. Images were acquired at 0.2- μm intervals along the z-axis and were subjected to volume deconvolution using SoftWoRx software.

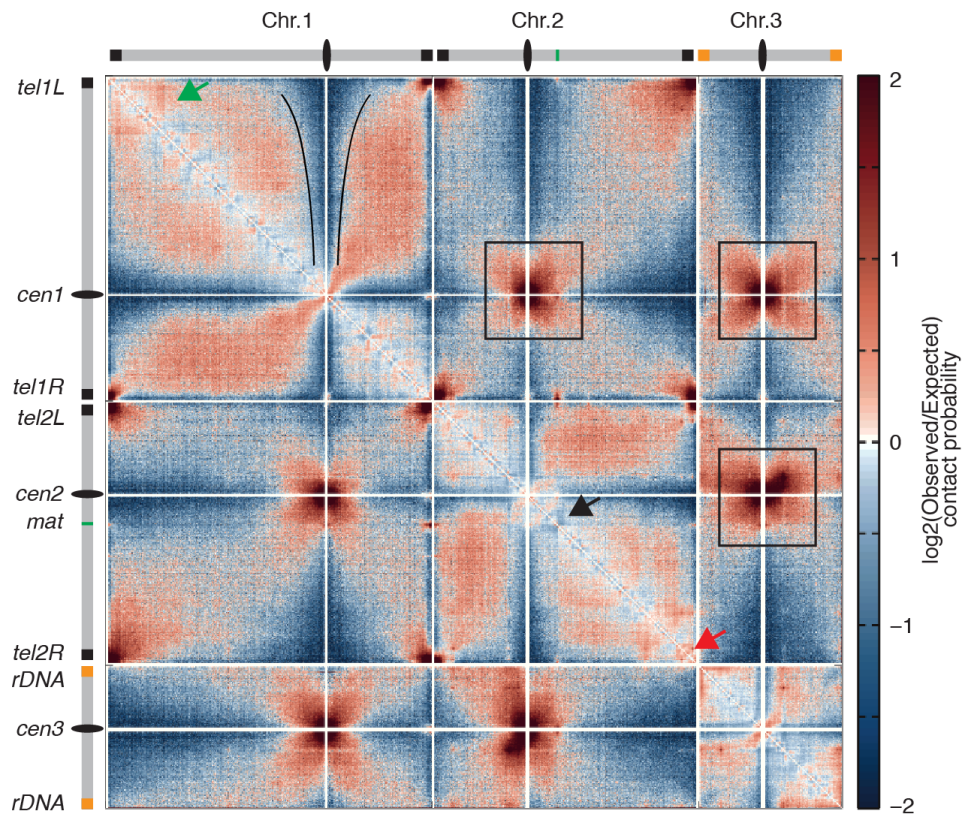
SPB or Cnp3–lacO distances live imaging. Distances between the spindle pole body (Sad1–mCherry) or kinetochore protein CENP-C (Cnp3–Tomato) (gift from Y. Watanabe) and a lacO array at the *lys1* locus (~ 24 kb from central core 1)³⁴ were measured on G2 cells, displaying a single lacI–GFP dot, as follows. Cells were grown overnight until logarithmic phase in minimal medium EMM plus supplements at 30 °C and then mounted in 2% agarose pad. Cells were imaged on a Delta Vision Elite microscope (GE Healthcare) with a $\times 100$ 1.4 NA Plan Super Apochromat oil lens (Olympus). Thirty 0.2- μm z-sections were acquired and subsequently deconvolved using SoftWoRx 6.0 (GE Healthcare). Further image processing, including maximum intensity projections and measuring distances between mCherry/Tomato and GFP (centre-to-centre) was performed using ImageJ (National Institutes of Health).

Culture conditions for detecting genomic rearrangements. To detect rearrangements in *rad21-K1*, wild-type and mutant cells were cultured in YEA at 26 °C before shifting to 33 °C. DNA prepared from cells grown overnight was used to perform comparative genomic hybridization analyses as described below.

Comparative genomic hybridization. Comparative genomic hybridization analysis was performed using our custom Agilent microarray (4 \times 44 K format)¹². Genomic DNA from wild type or mutants was digested with AluI and RsaI. After complete digestion, mutant DNA was labelled with Cy-5 dCTP (Amersham Biosciences) and wild-type DNA was labelled with Cy-3 dCTP (Amersham Biosciences) using the BioPrime Array CGH Genomic Labelling kit (Invitrogen). Equal amounts of labelled DNA (1.5 μg) were competitively hybridized onto the microarray. Pre-hybridization, probe hybridization, washing and drying steps for arrays were performed as for ChIP-chip experiments¹². Arrays were scanned using an Agilent scanner (Agilent) and analysed using Agilent Feature Extraction (Agilent). Signal intensity ratios between Cy5 (mutant) and Cy3 (Wild type) were calculated from rProcessedSignal and gProcessedSignal values according to Agilent Feature Extraction. The \log_2 transformed Cy5/Cy3 ratio is plotted along the chromosome.

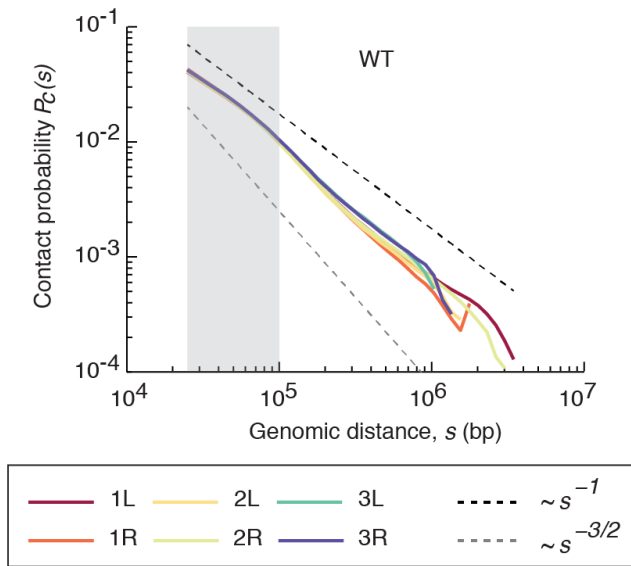
Chromatin immunoprecipitation and expression profiling. Chromatin immunoprecipitation was performed as previously described using anti-GFP (ab290 Abcam)¹². Expression profiling was carried out according to a protocol described previously³⁵.

29. Belton, J. M. *et al.* Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* **58**, 268–276 (2012).
30. Imakaev, M. *et al.* Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nature Methods* **9**, 999–1003 (2012).
31. Fudenberg, G. & Mirny, L. A. Higher-order chromatin structure: bridging physics and biology. *Curr. Opin. Genet. Dev.* **22**, 115–124 (2012).
32. Lantermann, A. B. *et al.* *Schizosaccharomyces pombe* genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of *Saccharomyces cerevisiae*. *Nature Struct. Mol. Biol.* **17**, 251–257 (2010).
33. Cam, H. P., Noma, K., Ebina, H., Levin, H. L. & Grewal, S. I. Host genome surveillance for retrotransposons by transposon-derived proteins. *Nature* **451**, 431–436 (2008).
34. Nabeshima, K. *et al.* Dynamics of centromeres during metaphase-anaphase transition in fission yeast: Dis1 is implicated in force balance in metaphase bipolar spindle. *Mol. Biol. Cell* **9**, 3211–3225 (1998).
35. Zofall, M. *et al.* Histone H2A.Z cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. *Nature* **461**, 419–422 (2009).

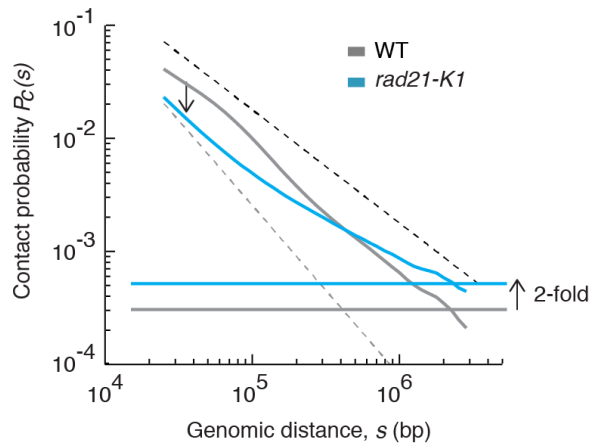


Extended Data Figure 1 | *S. pombe* chromosomes are partitioned into complex domains. The $\log_2(\text{observed/expected})$ contact probability maps for wild-type *S. pombe* chromosomes. The colour code indicates more (brown) or less (blue) interaction than expected depending on genomic distance (ranging from -2 to $+2$). These maps show an increasing extent of centromeric avoidance for regions along the arms of chromosomes (black outlines).

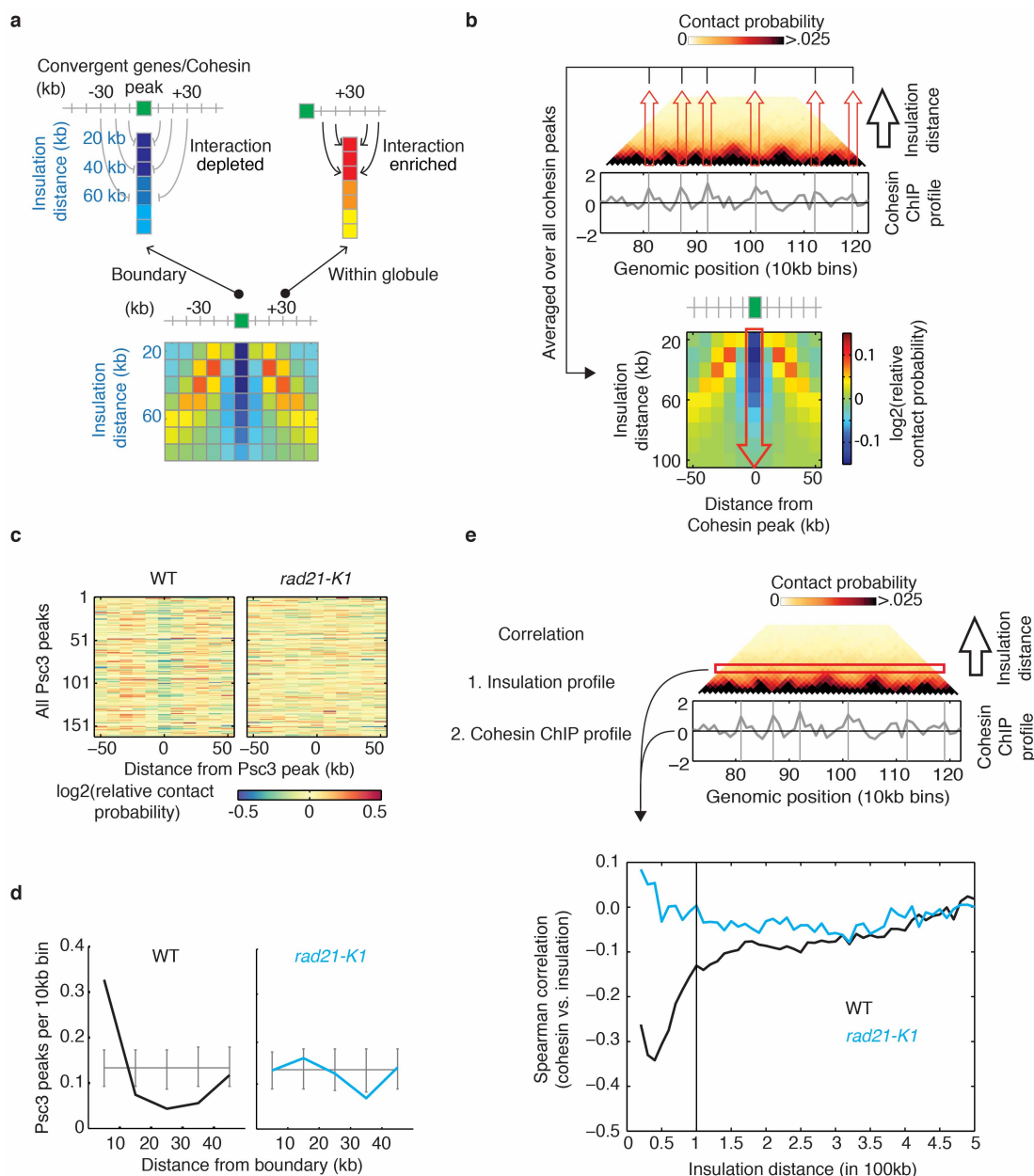
Boxes indicate strong cross-like patterns of centromere proximal arm-arm interactions. The left arm of chromosome 1 is segregated into compartments (green arrow). The mating type locus on chromosome 2 (black arrow) separates two adjacent domains. Subtelomere 2R (tel2R) is partitioned into strongly interacting domains (red arrow).



Extended Data Figure 2 | Contact probability as a function of genomic distance for different chromosomal arms. The decay of intra-arm contact probability as a function of genomic distance, plotted for each chromosome arm. All chromosome arms behave similarly in terms of their scaling. $P_c(s)$ decreases more slowly at short distances (grey shaded area). The black and grey dashed lines represent the slope for fractal globules (-1) and polymers in a melt ($-3/2$), respectively.

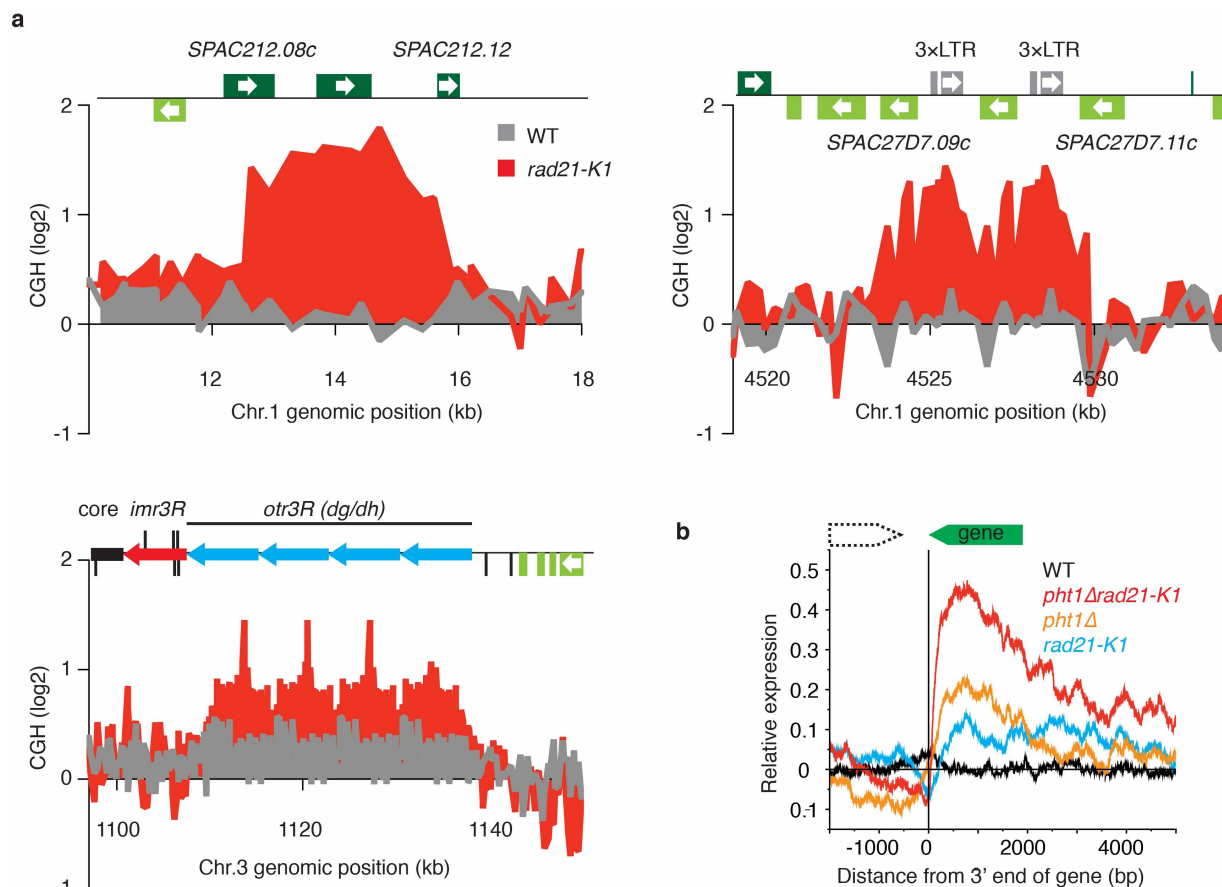


Extended Data Figure 3 | Contact probability as a function of genomic distance in wild-type and *rad21-K1*. Global decay of intra-arm contact probability as a function of genomic distance plotted for *rad21-K1* (blue) and wild type (grey). Average inter-chromosomal interactions (flat lines) are markedly increased (\sim twofold) in *rad21-K1* compared to wild type. Short-range contact probability decays rapidly in *rad21-K1*, and no longer decreases more rapidly after 100 kb, probably reflecting the loss of globules in *rad21-K1*. The black and grey dashed lines represent the slope for fractal globules (-1) and polymers in a melt ($-3/2$), respectively.



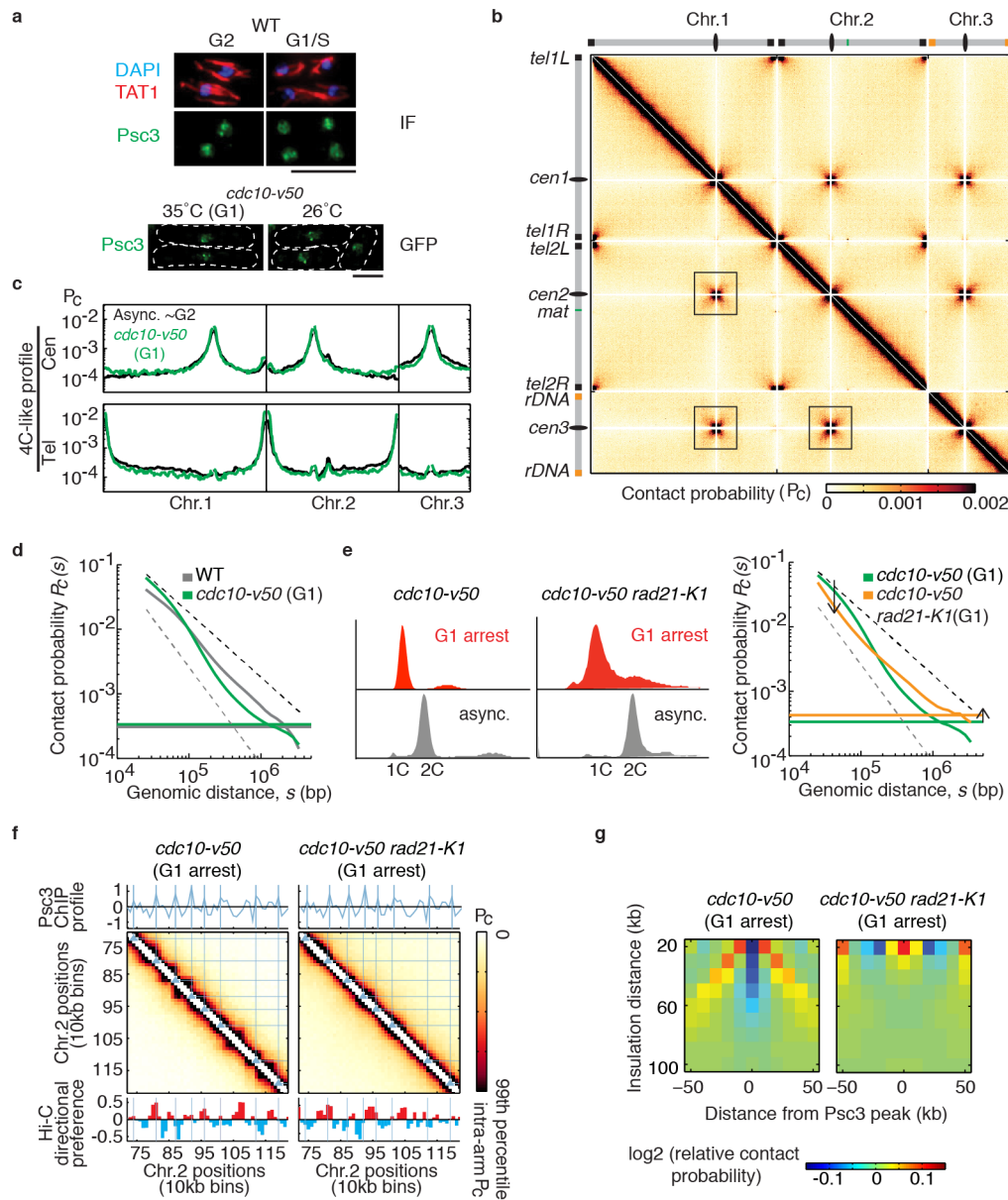
Extended Data Figure 4 | Insulation at globule boundaries. The degree to which a locus displays decreased contact frequency between, or is insulated from, regions separated by that locus can be directly quantified from the corrected Hi-C contact map. Here we use $R_j(s)$, the relative frequency of contacts occurring over a bin j at a distance s . Negative values of relative contact frequency, $R_j(s)$, are indicative of insulation at a given locus. $R_j(s)$ at a given distance s is calculated from a region within a rectangular band of a Hi-C contact map rotated by 45° . **a**, Diagram illustrating the concept of the insulation plot. At the location of the cohesin binding peak, interactions between two adjacent globules are less frequent (blue stripe). Within the globule domain, contact probability is high (red stripe). **b**, Relative contact probability around a cohesin peak as a function of insulation distance averaged over all cohesin peaks. Average insulation is examined by calculating the relative contact probability around cohesin peaks. Relative contact probability around the cohesin peak is depleted up to ~ 50 – 100 kb, indicative of insulation at peaks of local cohesin enrichment at these scales. **c**, Relative contact probability averaged

from 20–50 kb around positions of each cohesin peak (positions obtained in wild type were assayed in *rad21-K1*). **d**, Mean number of cohesin peaks as a function of distance from boundaries. Psc3 peaks are highly enriched at the boundary in wild type. **e**, The negative correlation between cohesin and relative contact frequency $R_j(s)$ in wild type indicates that not only is insulation observed at peaks of cohesin enrichment, but that the inverse relationship between the local enrichment of cohesin (Psc3) and the relative contact frequency holds genome-wide for data binned to 10 kb. This indicates that it is not just the presence or absence of a cohesin peak, but the local amount of cohesin protein in the chromatin fibre that may be important for boundary formation, as well as the strength of a given boundary. The negative correlation holds up to ~ 100 kb in wild type. In *rad21-K1*, however, there is no appreciable correlation with Psc3 at any distance. This indicates that there is no clear relationship between the distribution of cohesin and local chromatin organization in this mutant.



Extended Data Figure 5 | Genomic rearrangements and transcriptional dysregulation in *rad21-K1*. **a**, Microarray comparative genomic hybridization profile of *rad21-K1*. Genomic DNA isolated from *rad21-K1* and wild type was labelled with cy5-dCTP and cy3-dCTP, respectively. The $\log_2(\text{cy5/cy3}$ signal ratio) was plotted to detect copy number differences between the two strains. Several copy number gains were identified in *rad21-K1*. All changes were flanked by highly homologous sequences. *SPAC212.08c* and *SPAC212.12* share a 372-bp DNA stretch that shows 97% sequence similarity. *SPAC27D7.09c* and

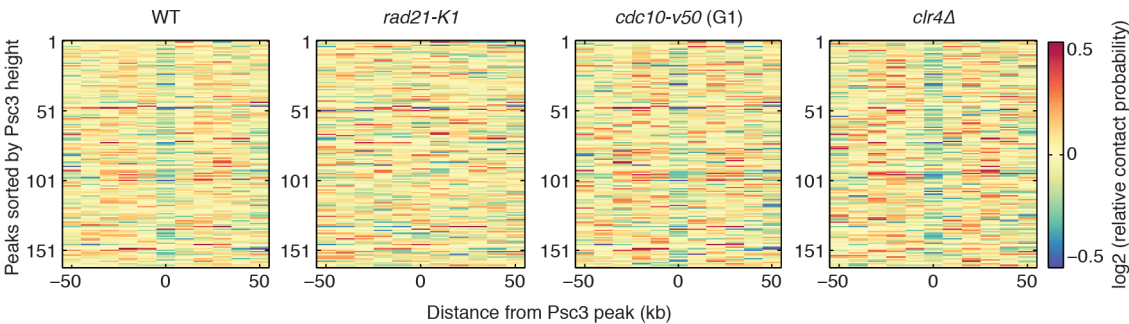
SPAC27D7.11c share a 560-bp DNA stretch that shows 88% sequence similarity. Pericentromeric heterochromatin contains a specific class of repeat elements, referred to as *dg/dh* repeats. **b**, Relative expression values (mutants/wild type) were plotted to detect read-through transcripts in the indicated strains. All pairs of convergent genes were aligned at the 3' end of the second gene in the pair. Note that *rad21-K1* cells show increased levels of read-through transcripts that were further enhanced in a *pht1Δ rad21-K1* double mutant lacking the histone variant H2A.Z known to prevent their accumulation³⁵.



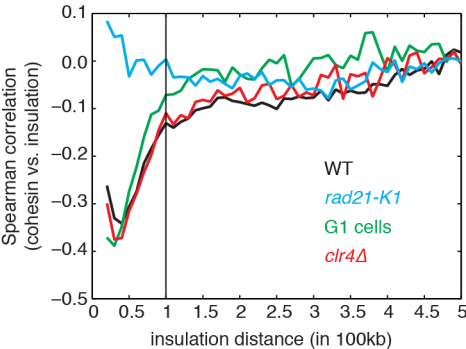
Extended Data Figure 6 | Hi-C analysis of G1-arrested cells. **a**, Cohesin (Psc3) localization was examined by immunofluorescence in asynchronous wild-type cells. Cell cycle stage was determined by tubulin staining (TAT1). Psc3 was detected in the nucleus in both G2 and G1/S phase cells (top). Psc3–GFP localization was examined in G1 arrested cells (*cdc10-v50*). Predominant nuclear staining and Psc3–GFP dots were detected in both asynchronous cells and G1-arrested cells (bottom). Scale bars, 5 μ m. **b**, All-by-all interaction heatmap for G1 cells. The inter-chromosomal cross-like pattern is more prominent in G1 cells than in asynchronous cells. **c**, 4C-like inter-chromosomal profiles for centromeres and telomeres. **d**, Global decay of intra-arm contact probability as a function of genomic distance in G1 cells (green) compared with wild type (grey); flat lines indicate average inter-chromosomal contact

probability. Slower decay of contact probability over short distances, followed by a more rapid decrease after 100 kb, was observed in G1-arrested cells. The black and grey dashed lines represent the slope for fractal globules (-1) and polymers in a melt ($-3/2$), respectively. **e**, FACS analysis of cell populations used for Hi-C (left). Global decay of intra-arm contact probability as a function of genomic distance in G1-arrested *rad21-K1* (orange) compared to G1 (green) cells is shown at right. **f**, Hi-C heatmaps of a segment of chromosome 2 for indicated samples overlaid with lines corresponding to cohesin peaks from the 10-kb binned cohesin (Psc3) profile. The Hi-C directional preference profile is shown below. Note the globules are not visible in G1-arrested *rad21-K1* (*cdc10-v50 rad21-K1*). **g**, Insulation plot around cohesin peak sites (detected in G1-arrested cells) for G1 and G1-arrested *rad21-K1*.

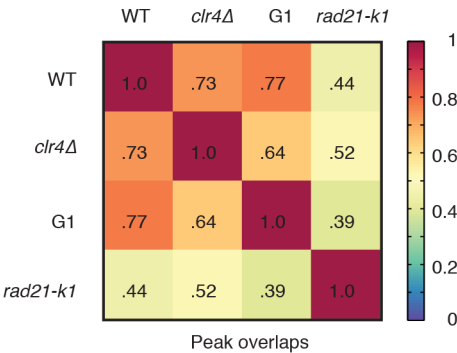
a



b

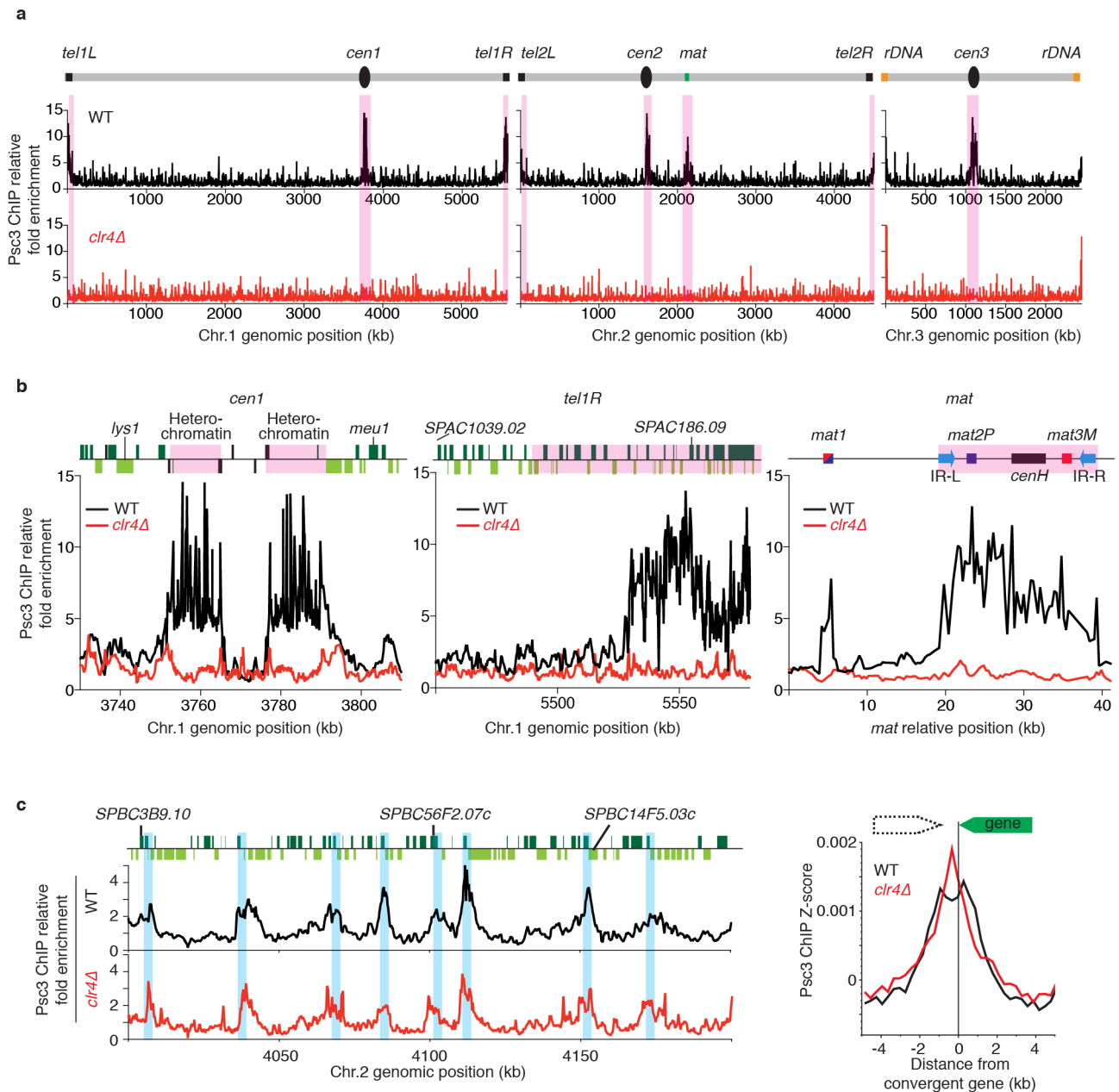


c



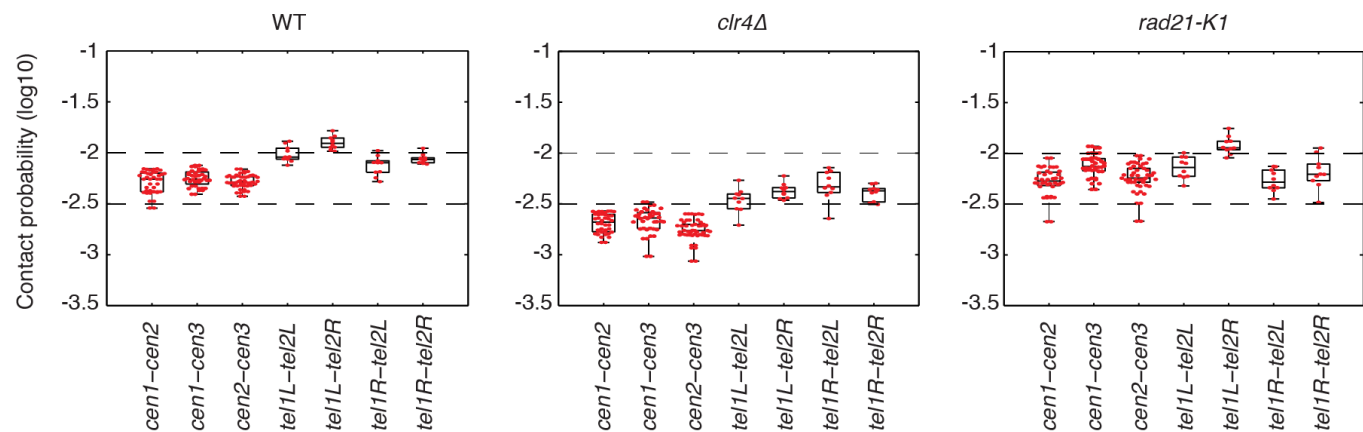
Extended Data Figure 7 | Globule boundaries in wild-type, *rad21-K1*, G1-arrested and *clr4Δ*. **a**, Insulation effect for all cohesin (Psc3) peaks. Relative contact probability was averaged over 20–50 kb at each Psc3 peak region (–50 kb to +50 kb region) and sorted by Psc3 peak height. Insulation effect at Psc3 peaks in wild type, G1 and *clr4Δ* holds genome-wide, and not only at a small subset of peaks. **b**, The correlation between Psc3 and relative contact probability $R_j(s)$ profile at a given distance s . The negative correlations hold up

to 100 kb in wild type, G1-arrested cells and *clr4Δ*. Psc3 enrichment and the relative contact probability profile in *rad21-K1* do not show an appreciable correlation at any distance. **c**, The fraction of overlapping boundaries (± 1 bin) for each Hi-C data set is shown. Boundaries in *clr4Δ* and G1-arrested cells show high overlap with wild-type boundaries. The top 100 strongest boundaries from each data set were examined for comparison.



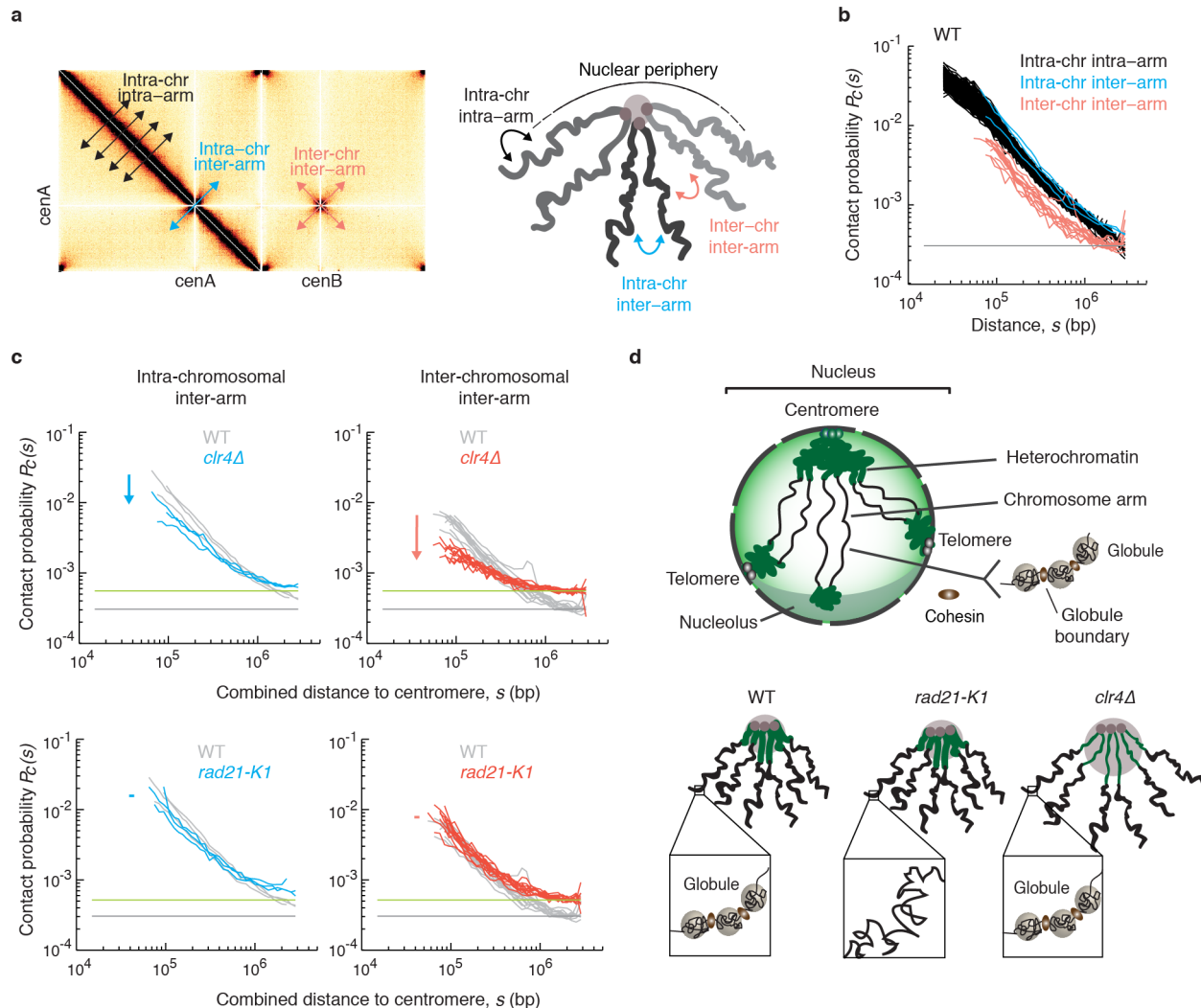
Extended Data Figure 8 | Cohesin localization in wild-type and *clr4Δ* strains as determined by ChIP-chip. **a**, Cohesin subunit Psc3 tagged with GFP (Psc3-GFP) is distributed broadly throughout the genome. Note that Psc3 localization in *clr4Δ* was specifically affected in heterochromatic regions, but not in chromosome arm regions. **b**, Psc3-GFP localization across pericentromere, subtelomere and mating type heterochromatic regions. Heterochromatic regions are highlighted. **c**, Psc3-GFP localization on

chromosome arm regions (left). Green bars represent open reading frames according to the 2007 *S. pombe* genome assembly. Cohesin enrichment sites are highlighted. Genome-wide profile of Psc3-GFP chromatin immunoprecipitation enrichment at convergent genes in wild type and *clr4Δ* (right). All pairs of convergent genes were aligned at the 3' end of the second gene (green box).



Extended Data Figure 9 | Box plots showing contact probabilities for centromere-centromere/telomere-telomere inter-chromosomal interactions in wild type, *clr4Δ* and *rad21-K1*. Box plots, overlaid with values for individual bin-pairs, showing contact probabilities for centromere-centromere/telomere-telomere inter-chromosomal interactions in wild type,

clr4Δ and *rad21-K1*. Whiskers span from minimum to maximum of each set of points, boxes show twenty-fifth, median and seventy-fifth percentiles. Centromere-centromere/telomere-telomere inter-chromosomal interactions are less frequent in *clr4Δ* as compared to wild type.



Extended Data Figure 10 | Effects of heterochromatin and cohesin on centromere proximal arm interactions, and a model showing their distinct effects on *S. pombe* genome organization. **a**, Centromere proximal regions considered for intra-chromosomal and inter-chromosomal arm interaction $P_c(s)$ are shown on a heatmap (left). Diagram showing intra- and inter-arm interactions within a chromosome, and inter-arm interactions between chromosomes in centromere proximal regions (right). Standard genomic distances were used to consider intra-arm contact probabilities. To consider inter-arm contacts within or between chromosomes, genomic distance was defined as the combined distance of two loci to the centromere, for regions at similar (<50 kb) distance from their respective centromeres. **b**, $P_c(s)$ plotted for inter-arm interactions within a chromosome and between different chromosomes in wild-type cells. Inter-chromosomal inter-arm (red) $P_c(s)$ falls below intra-chromosomal inter-arm (blue) and intra-arm (black), but have a similar rate of decay with distance. Note that inter-arm $P_c(s)$ starts at a larger genomic distance, since centromere proximal bins were removed at the stage of bin-level filtering due to their low coverage. **c**, Pairwise comparisons of inter-arm $P_c(s)$ of *rad21-K1* and *clr4Δ* with wild type. In *clr4Δ*, both inter-arm

$P_c(s)$ are shifted lower, most notably for inter-chromosomal inter-arm contact probability. In *rad21-K1*, inter-arm scaling near centromere is similar to wild type. **d**, Model showing distinct roles of heterochromatin and cohesin-dependent globules in overall chromosome organization. In wild-type cells, non-random organization consistent with a degree of chromosome territoriality was evident. These levels of organization may underlie genomic integrity, both independently and collectively, for example, by effectively preventing interaction between repetitive elements. The peripheral positioning of centromere and telomere clusters promotes a Rabl configuration. Heterochromatin reinforces this configuration by compacting centromere and telomere proximal regions, promoting strong interactions and aligning arms to facilitate proper genome architecture. A newly identified layer of globules bounded by high amounts of arm cohesin organizes chromosome arms. The formation of globules depends on arm cohesin. Unlike in wild type, globules on arms are disrupted in *rad21-K1*, whereas constraints at pericentromeric regions are maintained. In *clr4Δ*, pericentromeric regions are decompacted, but globules are not disrupted along arms.

R-loops induce repressive chromatin marks over mammalian gene terminators

Konstantina Skourti-Stathaki¹, Kinga Kamieniarz-Gdula¹ & Nicholas J. Proudfoot¹

The formation of R-loops is a natural consequence of the transcription process, caused by invasion of the DNA duplex by nascent transcripts. These structures have been considered rare transcriptional by-products with potentially harmful effects on genome integrity owing to the fragility of the displaced DNA coding strand¹. However, R-loops may also possess beneficial effects, as their widespread formation has been detected over CpG island promoters in human genes^{2,3}. Furthermore, we have previously shown that R-loops are particularly enriched over G-rich terminator elements. These facilitate RNA polymerase II (Pol II) pausing before efficient termination⁴. Here we reveal an unanticipated link between R-loops and RNA-interference-dependent H3K9me2 formation over pause-site termination regions in mammalian protein-coding genes. We show that R-loops induce antisense transcription over these pause elements, which in turn leads to the generation of double-stranded RNA and the recruitment of DICER, AGO1, AGO2 and the G9a histone lysine methyltransferase. Consequently, an H3K9me2 repressive mark is formed and heterochromatin protein 1 γ (HP1 γ) is recruited, which reinforces Pol II pausing before efficient transcriptional termination. We predict that R-loops promote a chromatin architecture that defines the termination region for a substantial subset of mammalian genes.

A connection between R-loops and heterochromatin formation was first made in fission yeast, in which removal of R-loops in centromeres caused a loss of heterochromatin structure⁵. An emerging theme is that heterochromatin and RNA interference (RNAi) machinery act broadly across the genome to regulate gene expression^{6,7}. Since processing of double-stranded (ds)RNA is the trigger for RNAi-dependent gene silencing, a source for the generation of dsRNA could be the hybridization of antisense transcripts with nascent pre-messenger RNA. We first detected localized antisense transcription over the termination region (pause site) of the human β -actin gene by reverse transcriptase-quantitative PCR (RT-qPCR) analysis (Fig. 1a). Next we tested for the formation of dsRNA by immunoprecipitation from whole HeLa cell extracts with the dsRNA-specific antibody, J2 (ref. 8). Selected RNA was analysed by strand-specific RT-qPCR. Positive signals for both sense and antisense transcripts were detected over 5' pause and pause regions (Fig. 1b, grey bars), suggesting that dsRNA is formed over these regions (Fig. 1b). dsRNA-specific V1, but not S1, nuclease treatments abolished sense and antisense signals, confirming dsRNA presence. DICER and AGO1 RNAi factors were also enriched over this region on the basis of chromatin immunoprecipitation (ChIP) analysis (Fig. 1c, d). Methylation of H3K9 is known to be the most conserved epigenetic mark associated with transcriptional silencing⁹. Since G9a and GLP are considered to be the major H3K9me1 and H3K9me2 histone lysine methyltransferases (HKMTs) of euchromatin^{10,11}, we performed ChIP analysis using anti-G9a antibody, which again showed G9a enrichment around the pause element (Fig. 1e). We also confirmed that H3K9me2 marks occur over the termination regions of the human β -actin gene (Fig. 1f and Extended Data Fig. 1a, b). H3K9me creates a binding site for the chromodomain of HP1 proteins (HP1 α , β and γ). We further show that HP1 γ is enriched over our heterochromatin terminator (Fig. 1g), consistent with the known association of HP1 γ with active genes^{12–15}. This suggests that HP1 γ acts as a

heterochromatin 'reader' over the human β -actin R-loop termination region. Previously, we showed using transfected gene constructs that R-loops are associated with termination regions, comprising a functional

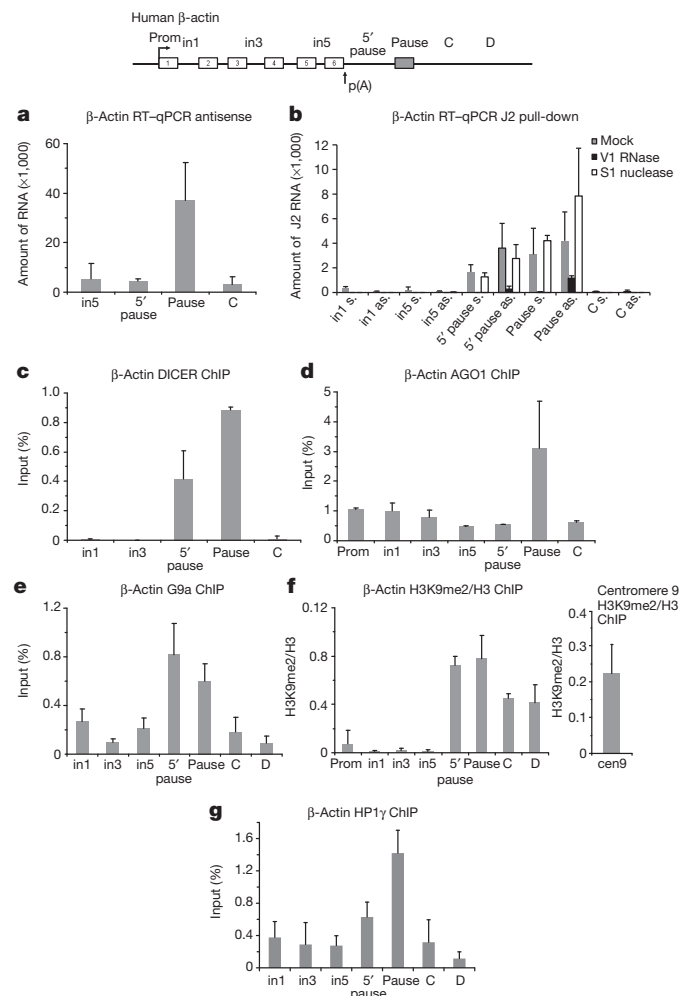


Figure 1 | The RNAi-dependent H3K9me2 repressive mark is formed over the human β -actin terminator in HeLa cells. **a**, RT-qPCR of β -actin antisense transcription. RT with region-specific forward primers. **b**, Sense and antisense transcripts levels determined by RT-qPCR from J2 immunoprecipitated dsRNA. Samples were untreated (grey bars), or treated with V1 RNase (black bars) or S1 nuclease (white bars). All RT-qPCR values are average \pm standard deviation (s.d.) from 3–4 biological repeats. **c–e, g**, ChIP analysis using DICER (**c**), AGO1 (**d**), G9a (**e**) and HP1 γ (**g**) antibodies. **f**, Ratio of H3K9me2 ChIP versus H3 on β -actin gene (left) and centromere 9 (right). ChIP values \pm s.d. from three biological repeats are shown. C and D indicate C and D termination probes, respectively. as., antisense; cen, centromere; in, intron; p(A), poly(A); prom, promoter; s., sense.

¹Sir William Dunn School of Pathology, South Parks Road, University of Oxford, Oxford OX1 3RE, UK.

poly(A) signal (PAS) and G-rich pause elements⁴. To investigate whether R-loops and the H3K9me2 mark are specific features of the pause-dependent termination mechanism, we employed cyclin B1 and akirin 1 genes that utilize alternative CoTC terminators¹⁶. DNA immunoprecipitation (DIP) and H3K9me2 ChIP analyses (Extended Data Fig. 2) showed no R-loop or H3K9me2 marks over these CoTC terminators, suggesting that such features are restricted to genes possessing pause-site terminators.

We investigated whether R-loops promote the recruitment of RNAi factors and H3K9me2 formation over the termination region of the β -actin gene by testing their sensitivity to RNase H1. Overexpression of this enzyme diminished R-loop levels over both the gene body (intron 1 amplicon) and pause regions (Fig. 2a). Remarkably, antisense RNA, DICER, G9a and HP1 γ occupancy were also diminished (Fig. 2b–e). To selectively remove the H3K9me2 repressive mark, we used the chemical inhibitor of G9a/GLP, BIX-01294 (BIX), which induces transient reduction of H3K9me2 in mammalian chromatin¹⁷. BIX treatment decreased the H3K9me2 signal over the 5' pause and pause regions as compared with non-treated cells (Fig. 2f and Extended Data Fig. 1c). Finally, we investigated whether R-loops are the consequence or cause of H3K9me2 occurrence. Notably, R-loop signals were unaffected by H3K9me2 reduction (Fig. 2g). This predicts that R-loops formed around the β -actin pause element trigger antisense transcription, assembly of the RNAi apparatus, the formation of an H3K9me2 repressive mark, and ultimately HP1 γ recruitment. We confirmed the correlation of R-loops and dsRNA with the H3K9me2 mark in single cells using immunofluorescence. We first demonstrated the nuclear localization of R-loops and dsRNA (Extended Data Fig. 3a). Ninety per cent of HeLa cell nuclei had R-loops and dsRNA in close proximity with H3K9me2 foci. Notably, 10% of either dsRNA or R-loop foci co-localized with H3K9me2 (Fig. 2h and Extended Data Fig. 3b). These data strongly suggest that, at a cellular level, R-loops are associated with gene silencing.

We next employed cell lines derived from mouse gene knockouts for *Ago2* and *G9a/Glp* to test their role in Pol II termination. We initially validated the data obtained with the human β -actin gene for its mouse homologue (Extended Data Figs 4–6). Notably, we observed that the repressive mark associated with the termination region of the mouse β -actin gene is specifically H3K9me2, and not H3K9me3 (Extended Data Fig. 4d). We then confirmed that Ago2, like Ago1, is specifically enriched at the termination region of the mouse β -actin gene and its recruitment is reduced to background levels in Ago2-knockout cells (Fig. 3a and Extended Data Fig. 4c). However, Ago1 recruitment is enhanced in Ago2-knockout cells, suggesting that Ago1 compensates for Ago2 depletion (Extended Data Fig. 4e). G9a (Fig. 3b) and H3K9me2 (Fig. 3c) ChIP analyses in Ago2-knockout cells showed a decrease in ChIP signals over the gene termination region, suggesting that the observed H3K9me2 mark is Ago2-dependent. However, the R-loop profile is Ago2-independent (Extended Data Fig. 7a), confirming that R-loops act upstream of the RNAi pathway. Similar results were obtained in *G9a/Glp* double-knockout cells (Extended Data Fig. 7c, d).

To investigate whether both R-loops and the H3K9me2 mark are needed for efficient transcriptional termination of the mouse β -actin gene, we performed Pol II ChIP in wild-type and Ago2-knockout mouse embryonic fibroblasts (MEFs) also overexpressing RNase H1. Pol II density increased, especially over termination probes C and D, indicative of a defect in transcriptional termination (Fig. 3d and Extended Data Fig. 7b). We also performed Br-UTP nuclear run-on (NRO) analysis (Fig. 3e) and detected significant enrichment of nascent read-through RNA signals over the termination region, relative to the gene body (intron 3 primer), in Ago2-knockout cells overexpressing RNase H1, as compared to wild-type cells. This suggests that R-loops and the H3K9me2 mark are both critical components of efficient pause-dependent Pol II termination. No nascent transcripts were detected over probes E and F, located 3.2 and 4 kilobases (kb) downstream of the PAS, suggesting that the effect of combined loss of the H3K9me2 mark and R-loops promotes read-through transcription up to 3 kb downstream of the PAS.

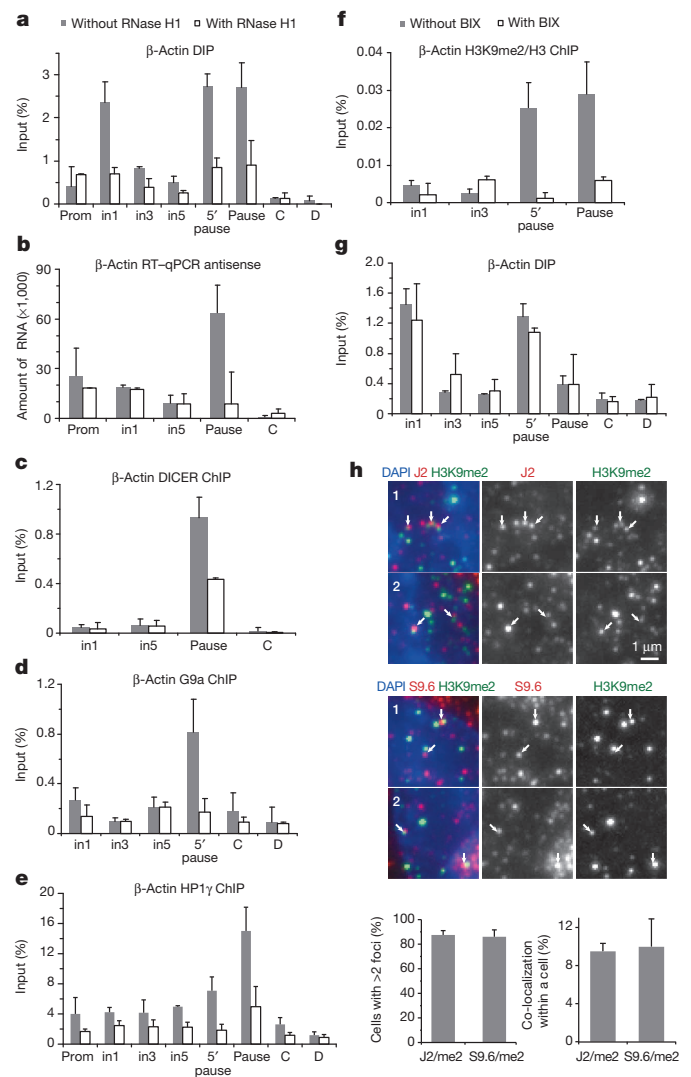


Figure 2 | Modulation of R-loop and G9a levels defines the mechanism of H3K9me2 formation on the human β -actin terminator. **a**, DIP with RNA:DNA hybrid antibody with or without RNase H1 overexpression. **b**, RT-qPCR with or without RNase H1 overexpression. **c–e**, ChIP analysis with or without RNase H1 overexpression using DICER, G9a or HP1 γ antibodies. **f**, H3K9me2 versus H3 ChIP values, with or without BIX treatment. **g**, DIP profile with or without BIX treatment. All ChIP and DIP values are \pm s.d. from three biological repeats. **h**, Nuclear immunofluorescence of H3K9me2 with dsRNA (J2; top) and R-loops (S9.6; bottom). Arrows denote foci in close proximity. DAPI, 4',6-diamidino-2-phenylindole. Numbers in panels denote images enlarged from indicated regions of whole cells shown in Extended Data Fig. 3b. Cell numbers with >2 J2/H3K9me2 and S9.6/H3K9me2 foci ($n = 100$) (bottom left graph). Co-localizing foci of J2 and S9.6 with H3K9me2 ($n = 1,000$), based on three independent experiments (bottom right graph). C and D indicate C and D termination probes, respectively. in, intron; prom, promoter.

We considered the possibility that RNAi-mediated heterochromatin formation induced by R-loop formation is a general termination mechanism, at least for a subset of genes. We performed a genomic meta-analysis of high-throughput sequencing of DNA derived from ChIP (ChIP-seq) data sets to look for the co-incidence of a paused elongating form of Pol II that is phosphorylated on Ser 2 of the carboxy-terminal domain (CTD) (PolIS2ph¹⁸), with HP1 γ enrichment¹⁵ within termination regions (Fig. 4a). We termed such regions of overlap pause-type termination (PTT) candidate regions. HP1 γ was previously implicated in transcriptional elongation^{12–15}. Indeed, 84% of the summits of HP1 γ peaks determined

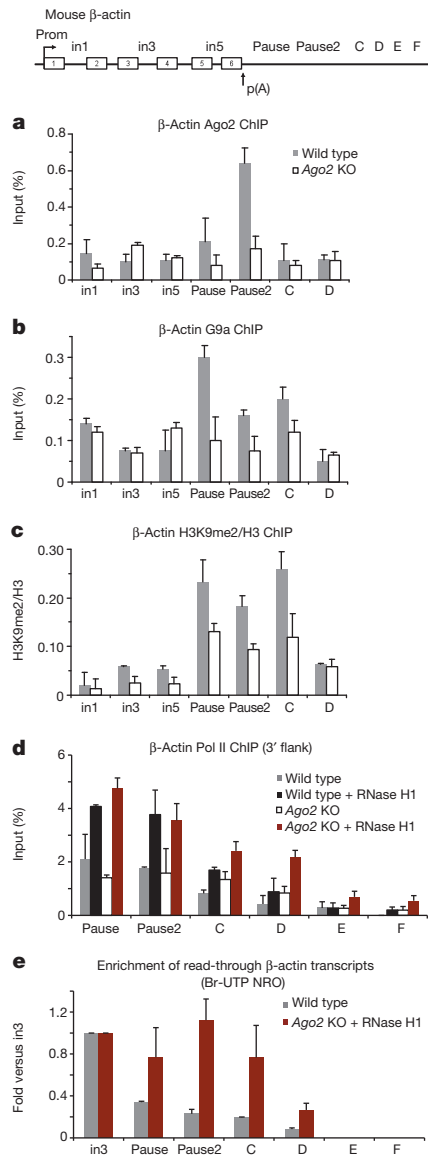


Figure 3 | Ago2-dependent H3K9me2 mark and R-loop formation promote efficient termination on mouse β -actin gene. **a**, **b**, ChIP in wild-type and Ago2-knockout (KO) MEFs using Ago2 (**a**) and G9a antibodies (**b**). **c**, Ratio of H3K9me2 versus H3 ChIP in wild-type and Ago2-knockout MEFs. **d**, Pol II ChIP with probes downstream of the PAS with extended y axis. Experiment was performed in wild-type (grey bars), wild-type overexpressing RNase H1 (black bars), Ago2-knockout (white bars) and Ago2-knockout overexpressing RNase H1 (red bars) MEFs. Full gene profile is in Extended Data Fig. 7b. All ChIP values are \pm s.d. from 3–4 biological repeats. **e**, Br-UTP NRO analysis in wild-type (grey bars) and Ago2-knockout MEFs overexpressing RNase H1 (red bars). Nascent Br-RNA over intron 3 probe is set as 1. Fold of enrichment of read-through transcripts for pause, pause2 and the C termination probe calculated relative to intron (in3) signal. D, E and F indicate D, E and F termination probes, respectively. Values are \pm s.d. from three biological repeats. p(A), poly(A); prom, promoter.

by ChIP-seq reside within gene bodies (Extended Data Fig. 8a)¹⁵. However, the highest-fold enrichment for HP1 γ relative to genomic annotation is over termination regions and the highest density of HP1 γ peak summits is detected downstream of PAS, genome wide (Extended Data Fig. 8a, b). Notably, 74% of HP1 γ enriched regions in termination regions overlap with PolIIIS2ph enrichment (Fig. 4b). PTT candidate regions show a statistically significant signal enrichment of the G9a ChIP hybridized to a genomic microarray (ChIP-chip)¹⁹, both compared to randomly sampled genomic regions of the same size as well as to non-PTT HP1 γ

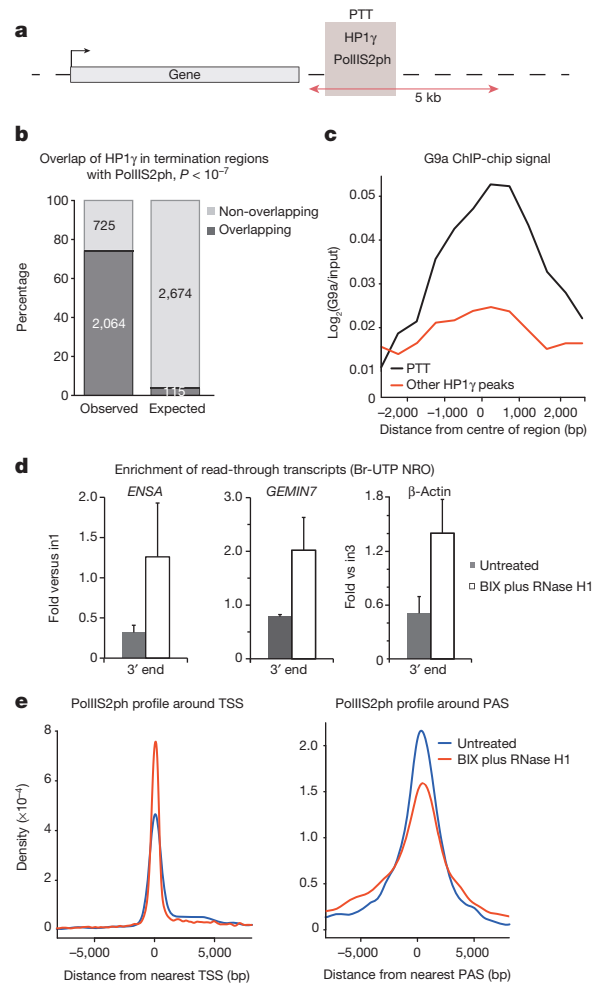


Figure 4 | HP1 γ , G9a and R-loops are globally associated with PTT regions. **a**, Diagram of PTT candidate regions. PTT candidate regions were defined as genomic intervals delineated by ChIP-seq peaks of PolIIIS2ph¹⁸ overlapping with ChIP-seq peaks of HP1 γ ¹⁵ within termination regions (5 kb downstream of human RefSeq genes, not overlapping with any downstream gene or promoter). **b**, Bar graph displaying the observed (2,064) and expected—based on random sampling (115)—overlap of HP1 γ terminator peaks with PolIIIS2ph peaks. **c**, G9a ChIP-chip profile over PTT candidate regions (black curve) and non-PTT-associated HP1 γ peaks (other HP1 γ peaks, red curve). **d**, Br-UTP NRO analysis with or without BIX treatment with RNase H1 overexpression on *ENSA*, *GEMIN7* and β -actin genes. Fold of enrichment of read-through transcripts over gene 3' end calculated relative to intronic signals (set as 1). Values \pm s.d. from three biological repeats. **e**, PolIIIS2ph ChIP-seq enrichment profiles for untreated (blue curve) and BIX plus RNase H1 overexpression (red curve) in 15 kb regions over the centre of the transcription start site (TSS) (left graph) and PAS (right graph).

peaks (Fig. 4c and Extended Data Fig. 8c), implicating H3K9 methyltransferase activity at these locations. To investigate whether PTT candidate regions are associated with R-loop formation, we compared the signal obtained by DNA:RNA-immunoprecipitation (DRIP) with that obtained by DRIP treated with RNase H1 (DRIPRH1 control) from previously published DRIP-seq data². PTT candidate regions show a significant enrichment of DRIP signal as compared to DRIPRH1 (Extended Data Fig. 8d, e), implying R-loop formation over these regions. We conclude that PTTs associated with R-loops, G9a and HP1 γ are widespread in the human genome.

Two genes, *ENSA* and *GEMIN7*, which show PolIIIS2ph pausing coincident with HP1 γ and DRIP-seq signal, were used to validate our genomic analysis. R-loops, antisense transcription, DICER, H3K9me2 and HP1 γ were observed over their termination regions (Extended Data Fig. 9), similar to the β -actin terminator. Finally, we performed Br-UTP NRO

analysis after BIX treatment and RNase H1 overexpression on these non-actin genes, showing that their termination requires R-loops and the H3K9me2 mark (Fig. 4d). The same effect was observed for the human β -actin gene, thus validating the data obtained in mouse β -actin (Fig. 3d, e).

Finally, to corroborate the role of R-loops and H3K9me2 on transcriptional termination genome wide, we performed ChIP-seq using an antibody against PolII S2ph on BIX-treated cells overexpressing RNase H1 (BIX RH1) and on untreated cells. We observe a decrease in PolII S2ph accumulation in the vicinity of PAS in the BIX RH1 sample versus the untreated sample (Fig. 4e, right). However, an increase in PolII S2ph accumulation around the transcription start site (TSS) is detected in the BIX RH1 sample versus the untreated condition (Fig. 4e, left). We then calculated the PolII S2ph pausing index in PTT candidate regions relative to gene bodies and observed that the BIX RH1 sample has a significantly lower value compared with the untreated sample ($P = 3.398 \times 10^{-16}$; Extended Data Fig. 8f). This implies that efficient pausing in these locations depends on the presence of R-loops and H3K9me2. By contrast, PolII S2ph pausing around the TSS concurrently increases in the BIX RH1 condition (Fig. 4e and Extended Data Fig. 8g), suggesting that termination and promoter pausing mechanisms are distinct. This is consistent with the specific enrichment of DICER, H3K9me2 and HP1 γ over gene 3' ends, but not promoter regions, of β -actin, *ENSA* and *GEMIN7* (Fig. 1 and Extended Data Fig. 9). Overall, we demonstrate that a termination mechanism mediated by Pol II pausing dependent on R-loop-induced heterochromatin is shared by a subset of human genes.

We reveal a molecular link between R-loop structures and the RNAi pathway. In particular, we have uncovered an unanticipated mechanism of regulated transcriptional termination through combined R-loops over pause-type 3' ends and epigenetic features. Our results predict a model for Pol II termination in which a G-rich sequence promotes R-loop formation, leading to heterochromatin establishment through the synthesis of localized dsRNA and the recruitment of RNAi factors (Extended Data Fig. 10). Previous studies^{14,20} support a functional association between histone marks, Pol II pausing and pre-mRNA processing. We now reveal that chromatin regulation at the level of transcriptional termination is mediated by the formation of R-loops. This raises the intriguing possibility that R-loops, a natural outcome of the transcription process, may more widely induce the formation of repressive chromatin marks to promote Pol II pausing.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 October 2013; accepted 22 August 2014.

Published online 5 October 2014.

1. Aguilera, A. & Garcia-Muse, T. R loops: from transcription byproducts to threats to genome stability. *Mol. Cell* **46**, 115–124 (2012).
2. Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I. & Chedin, F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol. Cell* **45**, 814–825 (2012).

3. Ginno, P. A., Lim, Y. W., Lott, P. L., Korf, I. & Chedin, F. GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *Genome Res.* **23**, 1590–1600 (2013).
4. Skourti-Stathaki, K., Proudfoot, N. J. & Gromak, N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell* **42**, 794–805 (2011).
5. Nakama, M., Kawakami, K., Kajitani, T., Urano, T. & Murakami, Y. DNA–RNA hybrid formation mediates RNAi-directed heterochromatin formation. *Genes Cells* **17**, 218–233 (2012).
6. Gullerova, M. & Proudfoot, N. J. Cohesin complex promotes transcriptional termination between convergent genes in *S. pombe*. *Cell* **132**, 983–995 (2008).
7. Zofall, M. *et al.* Histone H2A.Z cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. *Nature* **461**, 419–422 (2009).
8. Schönborn, J. *et al.* Monoclonal antibodies to double-stranded RNA as probes of RNA structure in crude nucleic acid extracts. *Nucleic Acids Res.* **19**, 2993–3000 (1991).
9. Jenuwein, T. The epigenetic magic of histone lysine methylation. *FEBS J.* **273**, 3121–3135 (2006).
10. Tachibana, M. *et al.* Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3–K9. *Genes Dev.* **19**, 815–826 (2005).
11. Rice, J. C. *et al.* Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol. Cell* **12**, 1591–1598 (2003).
12. Vakoc, C. R., Mandat, S. A., Olenchok, B. A. & Blobel, G. A. Histone H3 lysine 9 methylation and HP1 γ are associated with transcription elongation through mammalian chromatin. *Mol. Cell* **19**, 381–391 (2005).
13. Mateescu, B., Bourachot, B., Rachez, C., Ogryzko, V. & Muchardt, C. Regulation of an inducible promoter by an HP1 β –HP1 γ switch. *EMBO Rep.* **9**, 267–272 (2008).
14. Saint-André, V., Batsché, E., Rachez, C. & Muchardt, C. Histone H3 lysine 9 trimethylation and HP1 γ favor inclusion of alternative exons. *Nature Struct. Mol. Biol.* **18**, 337–344 (2011).
15. Smallwood, A. *et al.* CBX3 regulates efficient RNA processing genome-wide. *Genome Res.* **22**, 1426–1436 (2012).
16. Nojima, T., Dienstbier, M., Murphy, S., Proudfoot, N. J. & Dye, M. J. Definition of RNA polymerase II CoTC terminator elements in the human genome. *Cell Rep.* **3**, 1080–1092 (2013).
17. Kubicek, S. *et al.* Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. *Mol. Cell* **25**, 473–481 (2007).
18. Rosenbloom, K. R. *et al.* ENCODE data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res.* **41**, D56–D63 (2013).
19. Frieze, S., O'Geen, H., Blahnik, K. R., Jin, V. X. & Farnham, P. J. ZNF274 recruits the histone methyltransferase SETDB1 to the 3' ends of ZNF genes. *PLoS ONE* **5**, e15082 (2010).
20. Alló, M. *et al.* Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nature Struct. Mol. Biol.* **16**, 717–724 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by a Wellcome Trust Programme grant (091805/Z/10/Z) and a European Research Council Advanced grant (339270-polyloop) to N.J.P., and a Marie Curie Actions grant from EU FP7 (REA grant agreement, 327985) to K.K.-G. We are grateful to R. Nunes Bastos for valuable help with immunofluorescence/imaging analysis and figure formatting. We also thank Q. Sun for helpful discussions and advice. We thank R. J. Crouch for the gift of the GFP–RNase H1 expression plasmid, G. Hannon for the Ago2-knockout mouse cell line and Y. Shinkai for the G9a/Glp double-knockout mouse embryonic stem cells.

Author Contributions K.S.-S. performed all the molecular biology and imaging experiments. K.K.-G. performed the bioinformatics analysis. K.S.-S. and N.J.P. designed the experiments and wrote the manuscript.

Author Information Sequencing data for PolII S2ph ChIP and inputs from BIX RH1-treated and untreated HeLa cells have been deposited in the Gene Expression Omnibus under accession number GSE59878. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.J.P. (nicholas.proudfoot@path.ox.ac.uk).

METHODS

Molecular and cell biology techniques. Transfections of GFP-RNase H1 plasmid into human HeLa and mouse MEF cells were carried out as described previously⁴. Ago2-knockout and parental wild-type cells are MEFs. *G9a/Glp* double-knockout and their parental wild-type are mouse embryonic stem (mES) cells. Treatment with 10 μ M of BIX-01294 inhibitor (Sigma) was performed as described²⁰. Total RNA was isolated using TRIzol reagent (Invitrogen) and reverse transcribed with SuperScript III Reverse Transcriptase (Invitrogen) using gene-specific primers. J2 dsRNA pull-down was performed as described⁸. RT-qPCR levels are presented graphically as raw values $\times 1,000$. ChIP and genomic DIP analyses were carried out as before⁴. The following antibodies were used for ChIP: anti-H3K9me2 (Abcam), anti-H3K9me3 (Abcam), anti-H3 (Abcam), anti-DICER (13D6) (Abcam), anti-KMT1C/G9a (Abcam), anti-AGO1 (Millipore), anti-AGO2 (Abcam) and anti-Pol II (H-224) (Santa Cruz Biotechnology). S9.6 RNA:DNA-hybrid-specific antibody was used for DIP⁴. DNA oligonucleotide primers employed in these studies are listed in Supplementary Information Table 1.

Genomic analysis. Genomic interval processing, overlap calculations, statistical analysis and occupancy profiles were performed using custom scripts within the R/Bioconductor environment²¹.

The following publicly available human data sets were used: G9a ChIP-chip, Gene Expression Omnibus (GEO) accession GSE24480 (ref. 19); Pol II ChIP-seq, ENCODE-defined enriched regions (narrow peak) for HeLa-S3 using the phosphoS2 Pol II antibody ab5095, GEO accession GSE31477 (ref. 18); HP1 γ ChIP-seq, GEO accession GSE28115 (ref. 15), as well as R-loop locations delineated by DRIP-seq, Sequence Read Archive accession SRA048940.1 (ref. 2).

To obtain the HP1 γ peaks' coordinates, the mapped read coordinates were first lifted over from hg18 to hg19 using bedtools. The ChIP reads from two biological experiments—GSM699727 and GSM699729—were pooled, as well as their corresponding input reads, GSM699728 and GSM801616. The pooled ChIP and input raw reads were then used as input for peak calling using MACS2 with $q = 0.05$.

For genomic annotation, the coordinates of human hg19 RefSeq genes were downloaded from the UCSC table browser²² on 31 August 2012, and are synonymous with gene bodies throughout the manuscript. Promoter regions were defined as regions 1 kb upstream of RefSeq genes, excluding parts of intervals overlapping with any gene body. Terminator regions were defined as regions 5 kb downstream of genes, excluding those showing any overlap with gene bodies or promoters. As the only exception, for more accurate HP1 γ peak summit annotation (Extended Data Fig. 8a), a less stringent filtering approach for termination regions was applied: in case of partial overlap of the 5 kb downstream region with a gene body or promoter, the non-overlapping sequence interval was retained.

To calculate the P value of the overlap between terminator HP1 γ peaks and PolIIIS2ph peaks (Fig. 4b), 10^7 random data sets were generated based on PolIIIS2ph peaks and none of them had the same number or more overlaps than the original data set, hence P value $< 10^{-7}$. The average number of overlaps in the random data sets was 115.

PTT candidate regions were computed as genomic intervals corresponding to those ENCODE-defined PolIIIS2ph peaks or their fragments that reside within termination regions, and show a minimum 1-bp overlap with a HP1 γ peak¹⁵.

To obtain G9a, DRIP and HP1 γ occupancy profiles, their distance to the feature of interest was computed, and only retained if < 5 kb. Plotted is the frequency in 500-bp bins. In case of G9a ChIP-chip data the average \log_2 (G9a/input) signal was computed in 500-bp bins and subject to filtering using the moving average over six bins before plotting.

Deep sequencing of PolIIIS2ph ChIP and input from BIX RH1-treated and untreated HeLa cells was performed by the EMBL Genomics Core Facility. The two ChIP and two input samples were multiplexed, using NEBNext ChIP-Seq master-mix kit to prepare the libraries. The samples were sequenced on a 50-bp single-end run on three lanes using the Illumina HiSeq 2000 platform. Alignment of the sequenced tags to the hg19 human genome was performed using the CASAVA pipeline 1.8.2, ELAND parameters were: unique matches, 32 base seed, 2 mismatches allowed. This yielded a total of 115,813,632 reads uniquely aligned to hg19 for the untreated IP sample, 126,051,048 for the BIX RH1-treated sample, 127,749,851 for the untreated input and 113,690,799 for the BIX RH1-treated input.

Peak calling was performed on the IP samples versus their input controls using MACS2 with the parameters: $-q$ 0.05— $-nomodel$ — $-shiftsize$ 100. This procedure delineated 2,046 PolIIIS2ph peaks (enriched regions) for the untreated sample and 7,712

peaks for the BIX RH1-treated sample. We noted that the BIX RH1 treatment resulted in a higher overall PolIIIS2ph enrichment, presumably reflecting a globally more open chromatin environment following the treatment. Therefore, to avoid potential bias we chose to base our analysis on pausing indices relative to gene body signal (see later).

To obtain the PolIIIS2ph enrichment profiles over the TSS and PAS (Fig. 4e), the distance of the PolIIIS2ph peaks to the nearest TSS (or PAS) was computed, retaining only distances < 10 kb away from the feature of interest, which were then subject to kernel density estimation using the Gaussian smoothing kernel and plotted.

We defined the PTT PolIIIS2ph pausing index as a ratio of the normalized read density in PTT candidate regions to the normalized read density in its corresponding gene body. In more detail, we first re-computed the PTT candidate regions using the PolIIIS2ph peaks found in the untreated sample in place of the ENCODE data-set-derived peaks, and their corresponding gene body coordinates were extracted. The IP and input read number overlapping with each PTT and its corresponding gene body were counted and the IP and input reads per kilobase per million mapped reads (RPKM) read density for each region were computed as follows: $RPKM = (\text{number of reads overlapping with region}) / (\text{length of region in kb}) / (\text{million mapped reads})$. The RPKM value for input reads was then subtracted from the RPKM value for the IP reads for each region to yield the final normalized read density (NRD) for each region. Genes with low PolIIIS2ph NRD over their gene body ($NRD < 0.1$) were considered inactive and excluded from downstream analysis. The pausing index (PI) of each PPT/gene body pair was then computed as: $PI_{PPT} = NRD_{PPT} / NRD_{\text{gene body}}$.

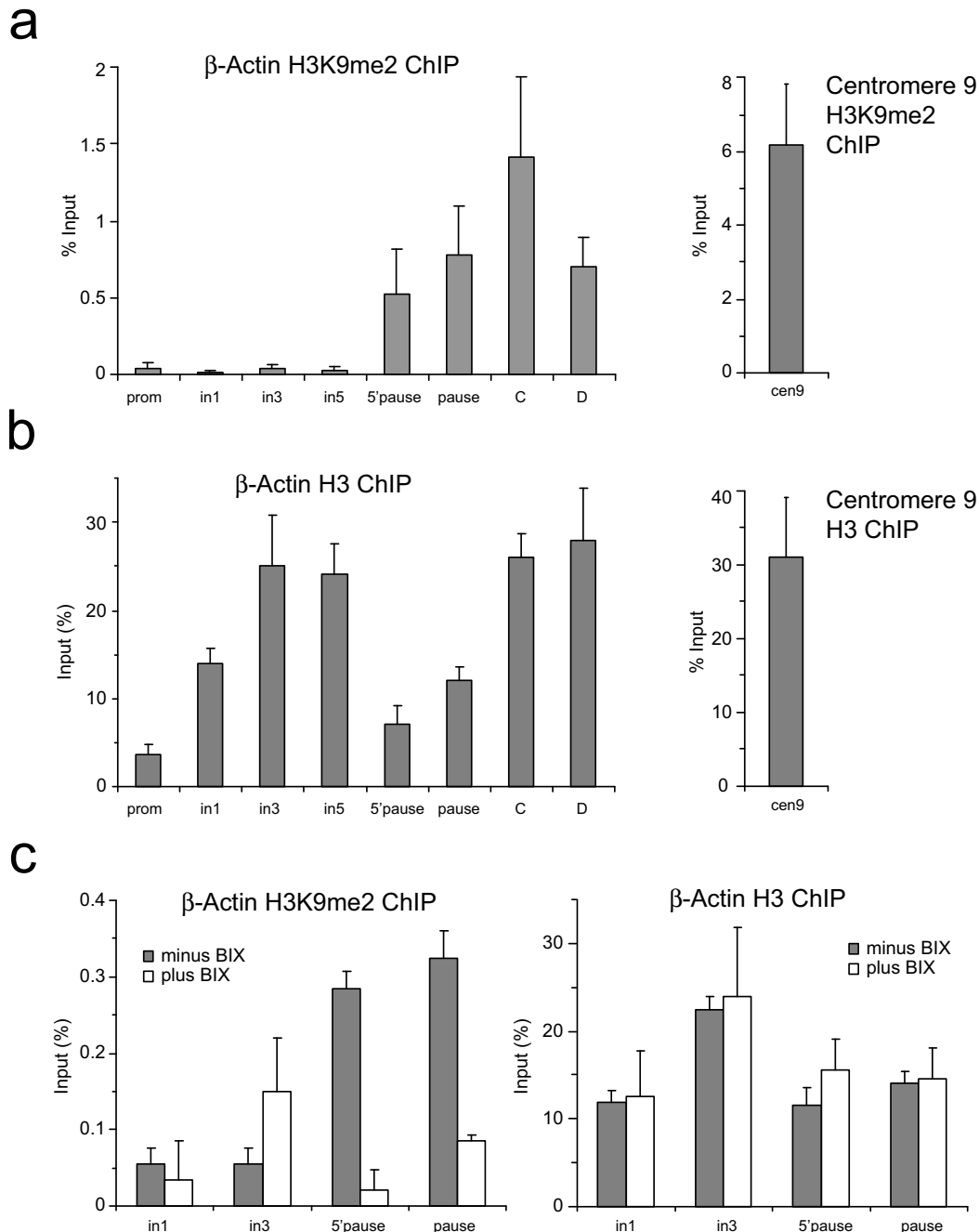
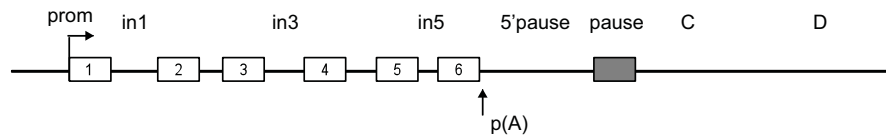
The pausing index for regions surrounding the TSSs of the PTT-linked genes by ± 1 kb was computed analogously: $PI_{TSS} = NRD_{TSS \pm 1 \text{ kb}} / NRD_{\text{gene body}}$. These computations were done in parallel for the BIX RH1-treated and untreated sample, and finally the distribution of the fold change in PolIIIS2ph pausing index between the BIX RH1-treated and untreated samples was calculated.

For statistical tests, since the data in Extended Data Fig. 8c, e–g did not conform to a normal distribution, non-parametric tests were employed: Wilcoxon signed-rank test for the paired samples in Extended Data Fig. 8e, f and Wilcoxon Mann-Whitney for the unpaired samples in Extended Data Fig. 8c, g. In all cases two-sided tests were applied.

Immunofluorescence and imaging analysis. Fixed cell samples were prepared and imaged exactly as described²³. In summary, cells grown on coverslips were fixed with 2 ml of ice-cold methanol or 3% paraformaldehyde in PBS for 15 min. Cells were quenched with 2 ml of 50 mM NH_4Cl in PBS for 10 min. Coverslips were washed three times in 2 ml PBS before permeabilization in 0.2% Triton X-100 for 5 min. In all cases primary and secondary antibody staining was performed in PBS for 60 min at room temperature. S9.6 antibody was used in 1:250 dilution, whereas commercial H3K9me2 (Cell Signaling) and J2 (Scicon) antibodies were used as directed by the manufacturers. 4',6-Diamidino-2-phenylindole (DAPI) was added to the secondary antibody staining solution at $0.3 \mu\text{g ml}^{-1}$. Coverslips were mounted in Mowiol 4-88 mounting medium (EMD Millipore). Fixed samples on glass slides were imaged using a $\times 60/\text{NA } 1.35$ oil immersion objective on an upright microscope (BX61; Olympus) with filtersets for DAPI, GFP/Alexa Fluor 488, 555, 568, and 647 (Chroma Technology), a CoolSNAP HQ2 camera (Roper Scientific) and MetaMorph 7.5 imaging software (Molecular Dynamics). Co-localization foci were measured as foci < 200 nm apart.

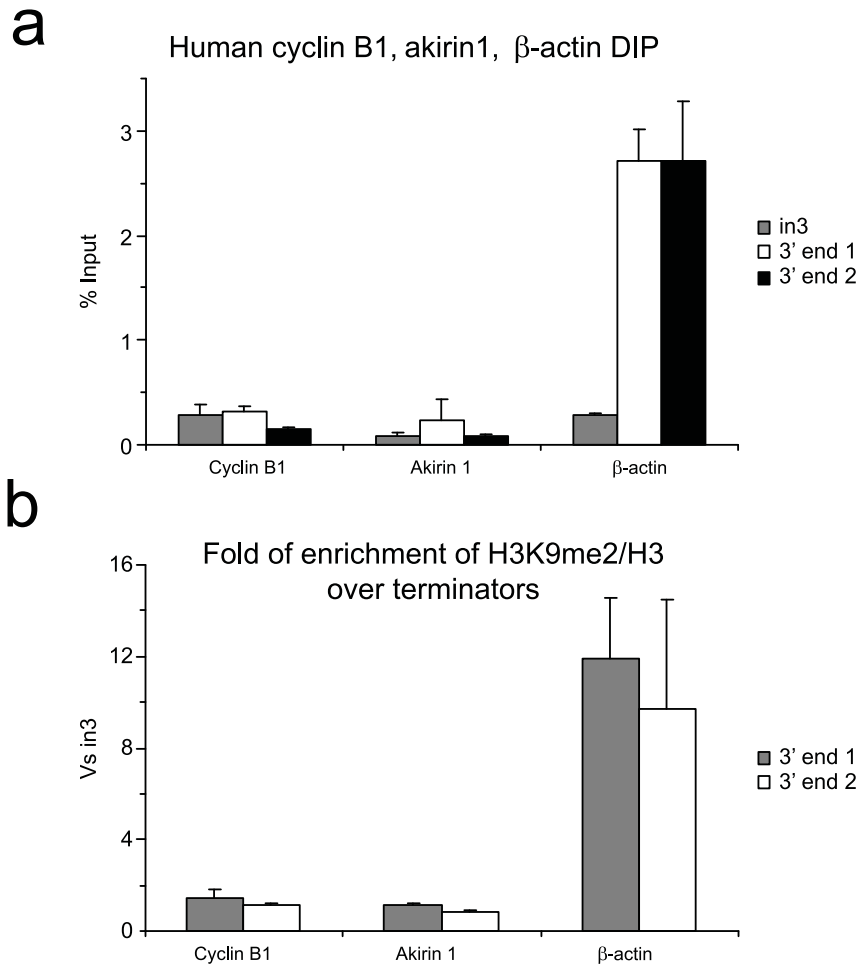
J2 dsRNA pull-down. J2 antibody (Scicon, 10010200, diluted to $0.1 \mu\text{g per } 1 \mu\text{g}$ of chromatin) was incubated with total cell extracts for 1.5 h on a rotating wheel at 4°C . Protein G-agarose beads (Millipore) were then added for an additional 1.5 h. dsRNA was then isolated from washed beads using the TRIzol reagent (Invitrogen) and analysed by RT-qPCR for sense and antisense transcripts. Signals from immunoprecipitated samples were subtracted from signals arising from non-precipitated samples. V1 and S1 treatments were carried out for 2 h at 37°C after the dsRNA isolation.

- Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).
- Karolchik, D. *et al.* The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* **32**, D493–D496 (2004).
- Bastos, R. N., Penate, X., Bates, M., Hammond, D. & Barr, F. A. CYK4 inhibits Rac1-dependent PAK1 and ARHGEF7 effector pathways during cytokinesis. *J. Cell Biol.* **198**, 865–880 (2012).

Human β -actin

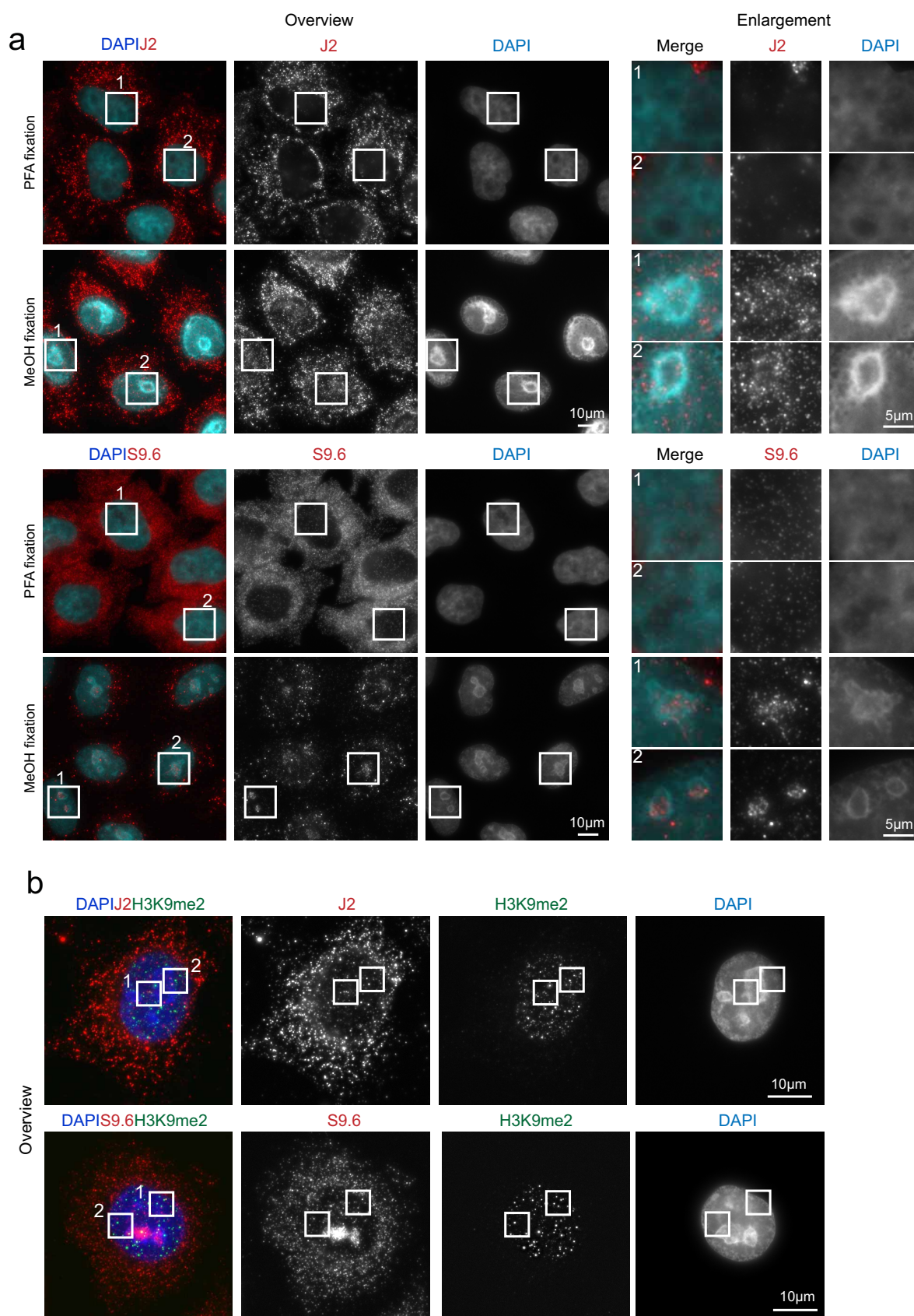
Extended Data Figure 1 | H3K9me2 and H3 levels over human β -actin gene.
a, Left, H3K9me2 ChIP on β -actin gene. Right, H3K9me2 ChIP analysis on human centromere 9 (positive control). **b**, Left, H3 ChIP on β -actin gene. Right, H3 ChIP analysis on human centromere 9. **c**, Left, H3K9me2 ChIP with or

without BIX treatment. Right, H3 ChIP with or without BIX treatment. ChIP values are \pm s.d. from three biological repeats. C and D indicate C and D termination probes, respectively. cen, centromere; in, intron; p(A), poly(A); prom, promoter.



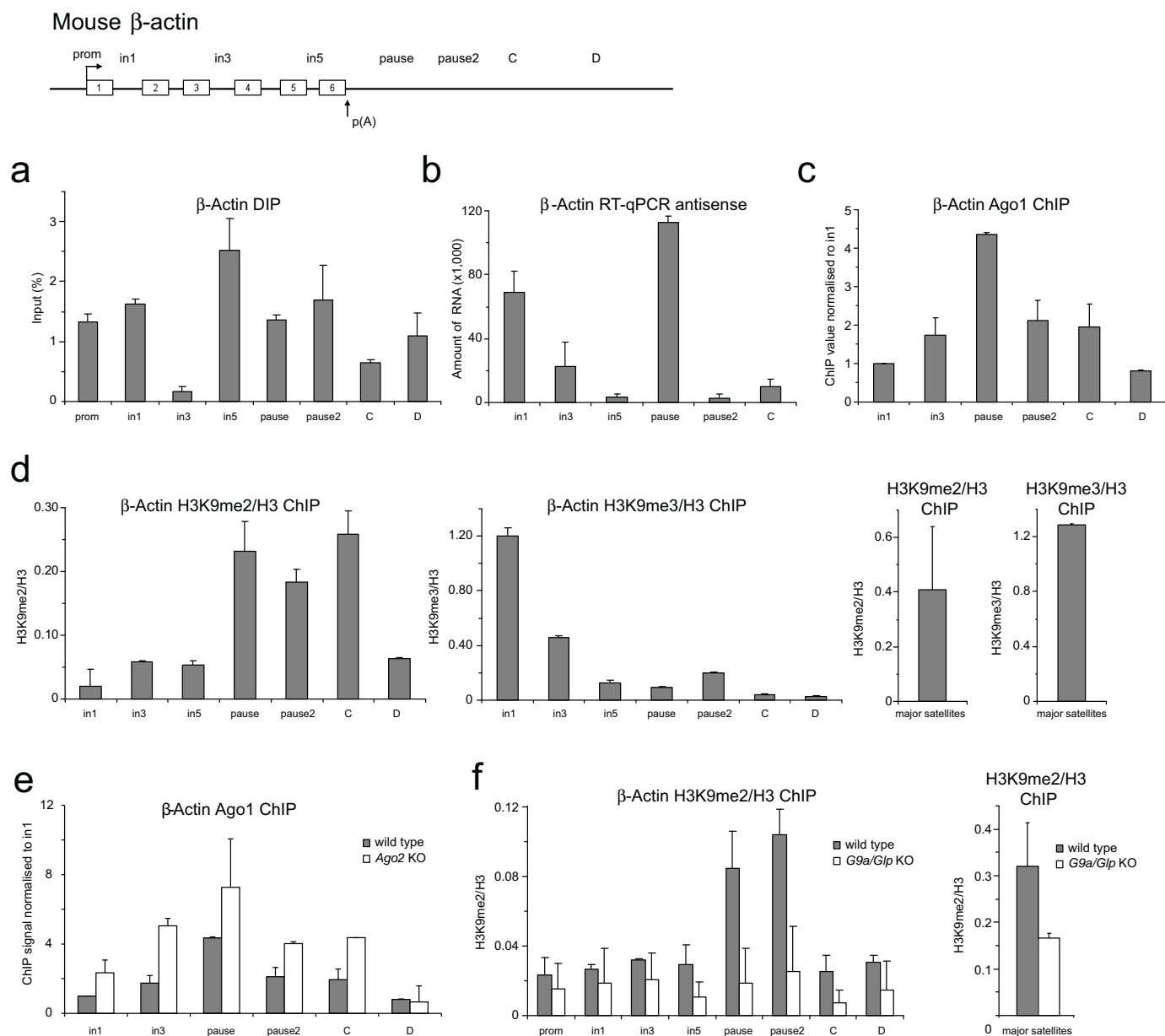
Extended Data Figure 2 | R-loops and the H3K9me2 repressive mark are not specifically enriched over the CoTC terminators of human cyclin B1 and akirin 1 genes. **a**, DIP on endogenous cyclin B1 and akirin 1 genes. No detection of R-loops was observed over their CoTC terminators. Human β -actin gene was used as a positive control. For cyclin B1 and akirin 1 genes, 3'-end 1 and 3'-end 2 amplicons amplify two different regions within the

CoTC terminator of each gene. 3'-end 1 and 3'-end 2 amplicons for the β -actin gene amplify the 5' pause and pause amplicons, respectively. **b**, Ratio of H3K9me2 signal over the 3' ends versus intron 3 signal in cyclin B1, akirin 1 and β -actin human genes. DIP and ChIP values are \pm s.d. from three biological repeats.



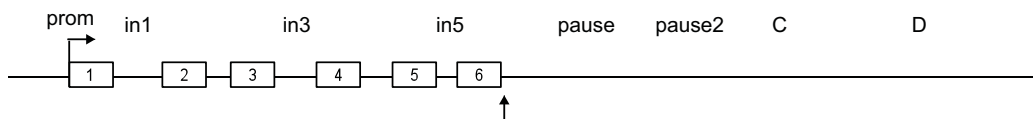
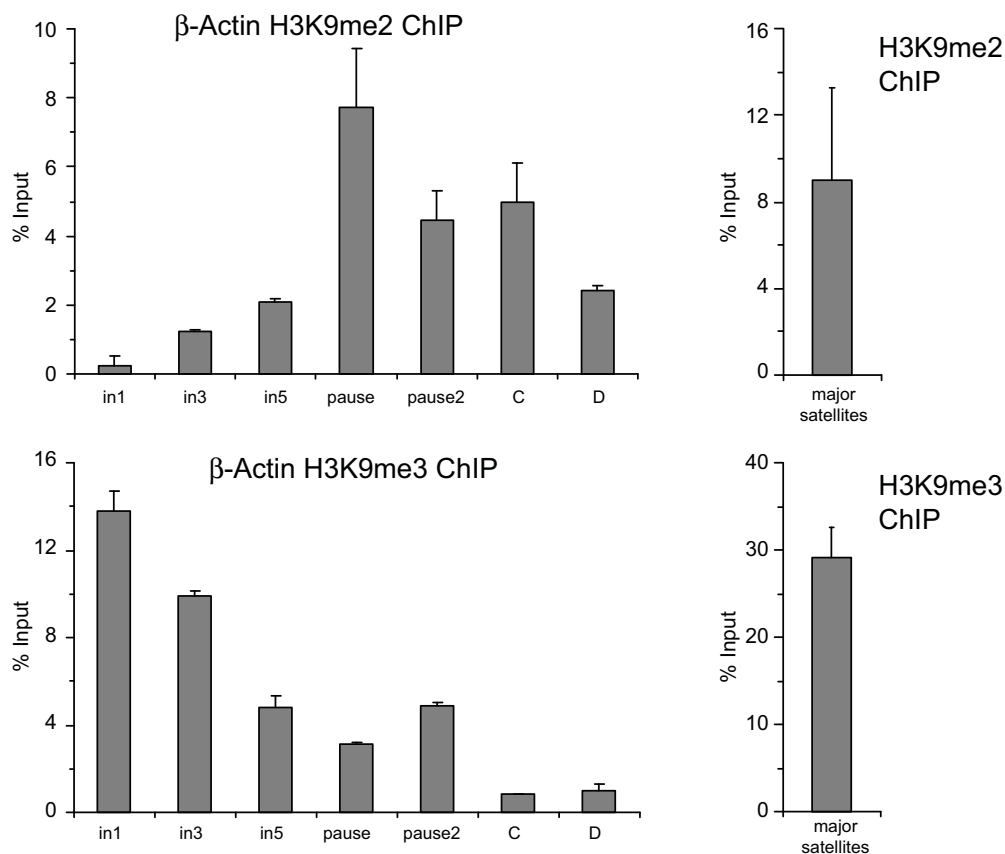
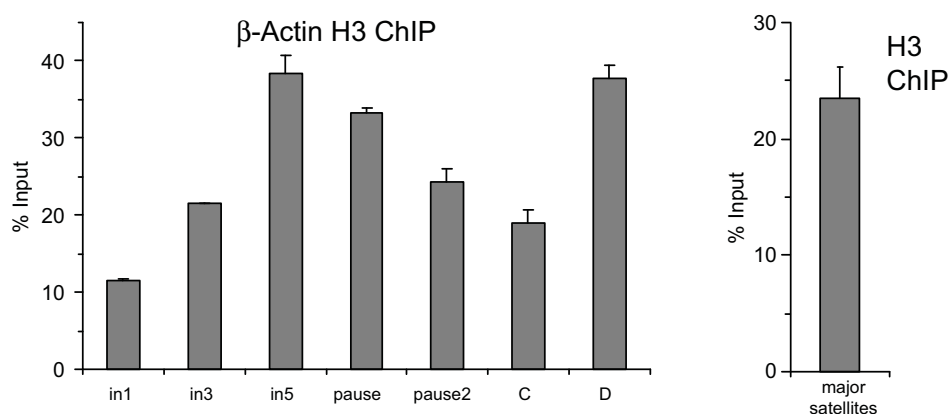
Extended Data Figure 3 | Cellular localization of R-loops, dsRNA and H3K9me2. **a**, Immunofluorescence imaging of dsRNA (J2 antibody) and R-loops (S9.6 antibody), using paraformaldehyde (PFA) and methanol (MeOH) as fixing reagents. Fixation with methanol allowed visualization of

R-loops and dsRNA in HeLa cell nuclei. Enlarged boxes (1 and 2) are shown in right panels. **b**, Whole-cell images showing immunofluorescence of H3K9me2 with dsRNA (J2; top) and R-loops (S9.6; bottom). Enlarged versions (1 and 2) are shown in Fig. 2h.



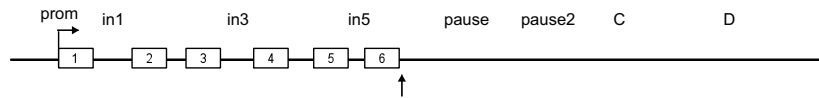
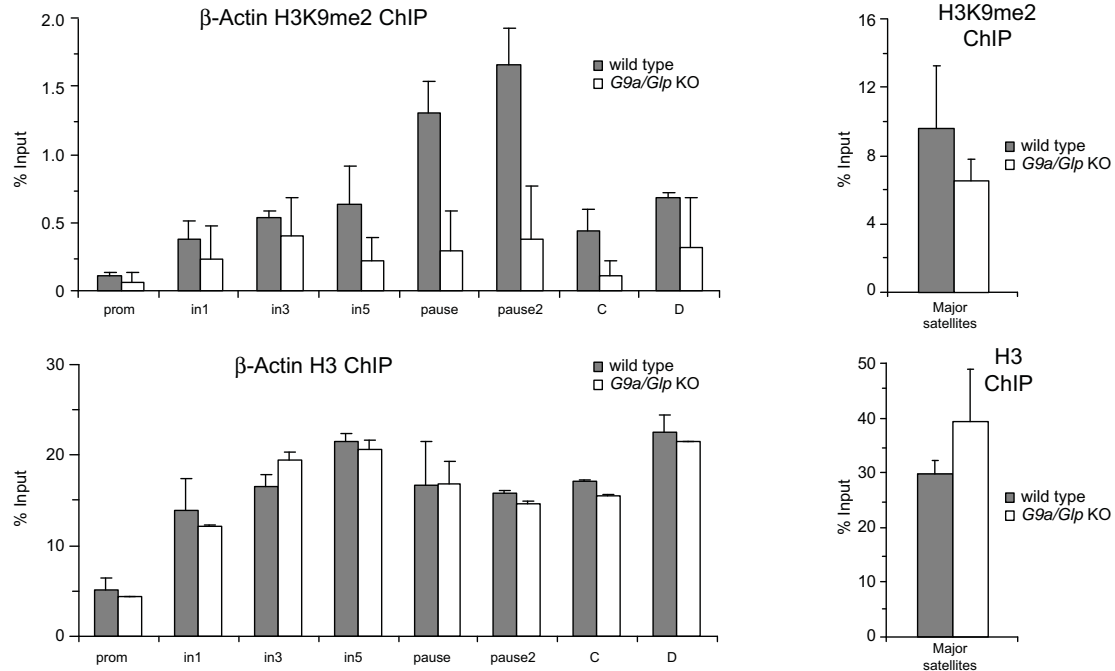
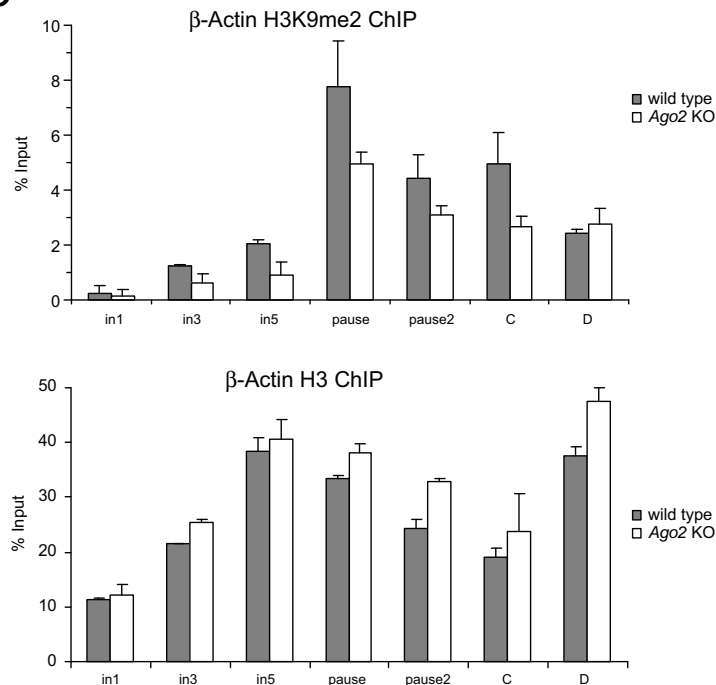
Extended Data Figure 4 | R-loops and RNAi promote the H3K9me2 mark over mouse β -actin terminator. **a**, DIP performed on mouse β -actin gene in MEFs. **b**, RT-qPCR of total RNA from MEF cells on β -actin gene to detect antisense transcripts with region-specific forward primers. Average RT-qPCR values are \pm s.d. from four biological repeats. **c**, Ago1 ChIP performed on mouse β -actin gene in MEFs. ChIP signal is normalized to intron 1 signal. **d**, Left, ratio of H3K9me2 ChIP signal versus H3 on mouse β -actin in MEFs. Middle, normalized H3K9me3 to total H3 levels. Right, ratio of H3K9me2 and

H3K9me3 signal versus H3 signal on major satellites in MEFs. **e**, Ago1 ChIP in wild-type (grey bars) and Ago2-knockout (KO) (white bars) cells. Ago1 recruitment over mouse β -actin is enhanced upon Ago2 depletion. **f**, Left, ratio of H3K9me2 ChIP signal versus total H3 on β -actin gene in wild-type and *G9a/Glp* double-knockout mouse embryonic stem cells. Right, H3K9me2/H3 ratio on the mouse major satellites in wild-type and *G9a/Glp* double-knockout cells. Average ChIP and DIP values are \pm s.d. from three biological repeats.

Mouse β -actin**a****b**

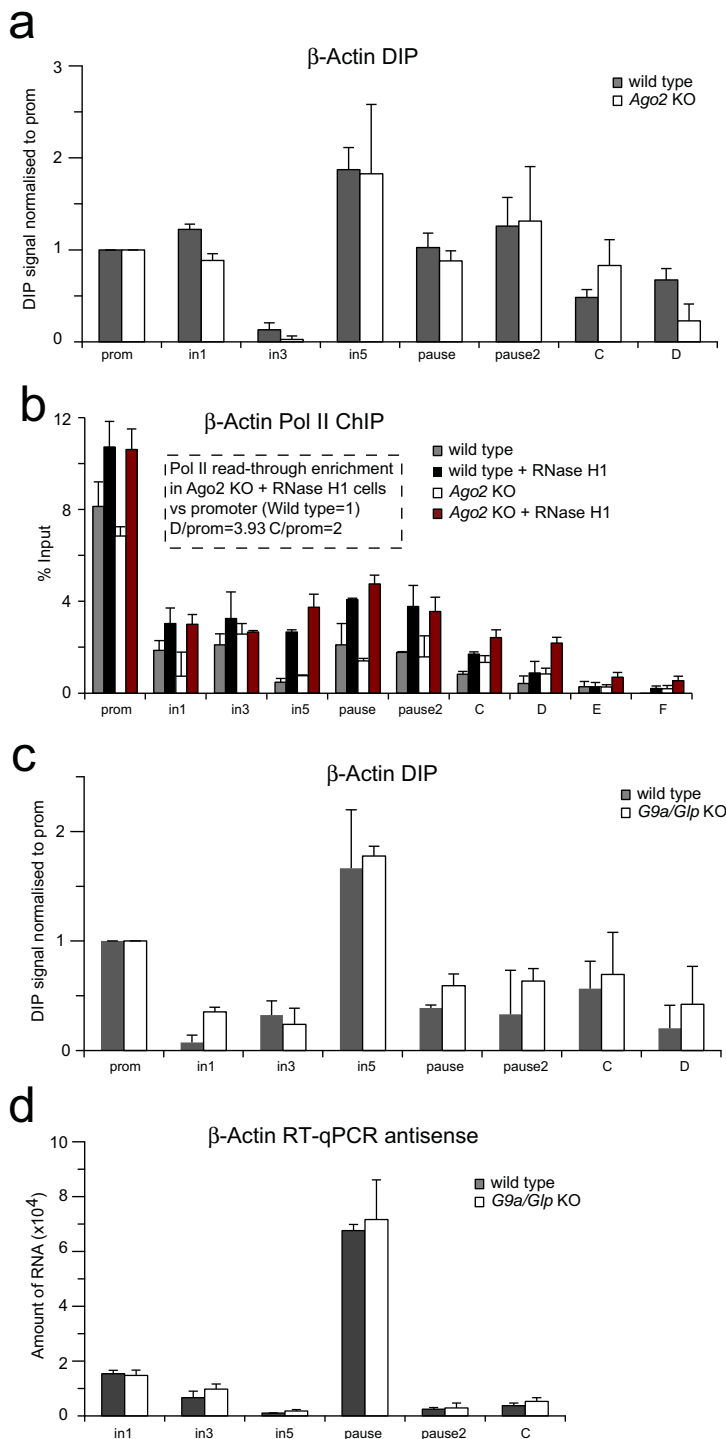
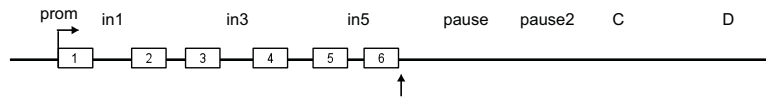
Extended Data Figure 5 | H3K9me2, H3K9me3 and H3 levels over the endogenous mouse β -actin gene. **a**, H3K9me2 and H3K9me3 ChIP on mouse β -actin gene in MEF cells. Right, H3K9me2 and H3K9me3 ChIP on mouse

major satellites (positive control). **b**, Total H3 ChIP on mouse β -actin gene. Major satellites were used as a positive control. ChIP values are \pm s.d. from three biological repeats.

Mouse β -actin**a****b**

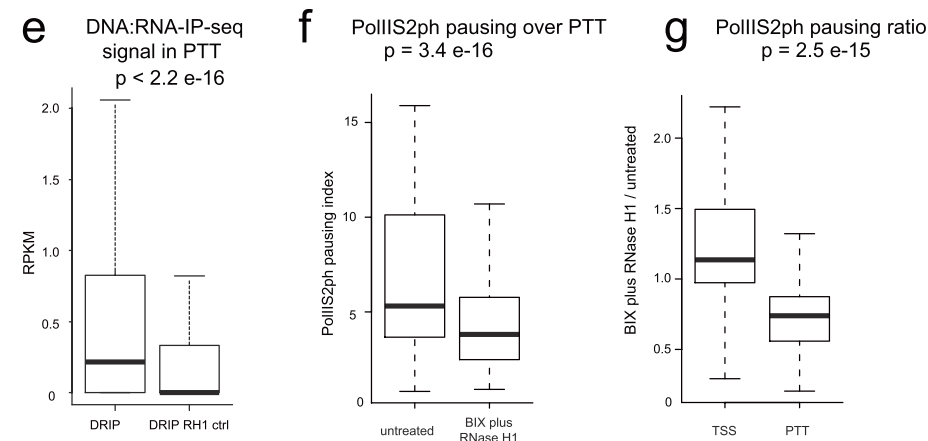
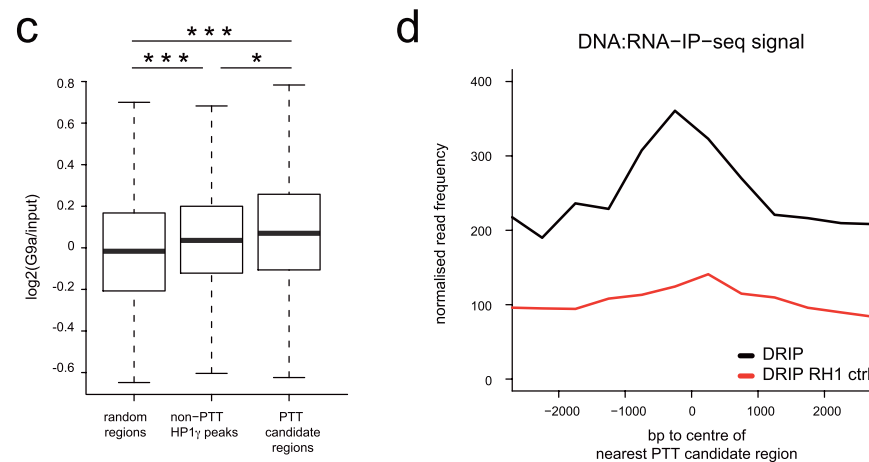
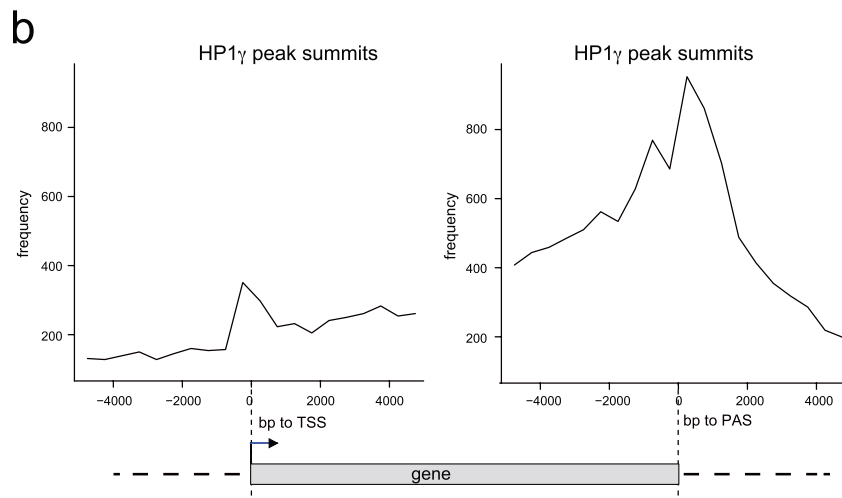
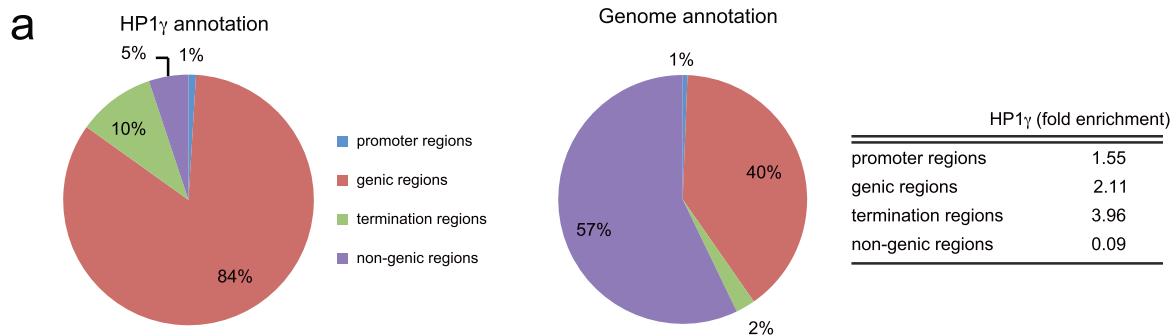
Extended Data Figure 6 | H3K9me2 and H3 levels over mouse β -actin gene in *G9a/Glp* double-knockout mouse embryonic stem cells and *Ago2*-knockout MEFs. a, Top and bottom, H3K9me2 and H3 ChIP performed on mouse β -actin gene in wild-type and *G9a/Glp* double-knockout embryonic stem cells. H3K9me2 occupancy depends on the presence of *G9a/Glp* HKMTs.

Right, H3K9me2 and H3 ChIP performed on mouse major satellites in wild-type and *G9a/Glp* double-knockout cells. b, ChIP analyses using H3K9me2 (top) and H3 (bottom) antibodies performed on mouse β -actin gene in wild-type and *Ago2*-knockout cells. ChIP values are \pm s.d. from three biological repeats.

Mouse β -actin

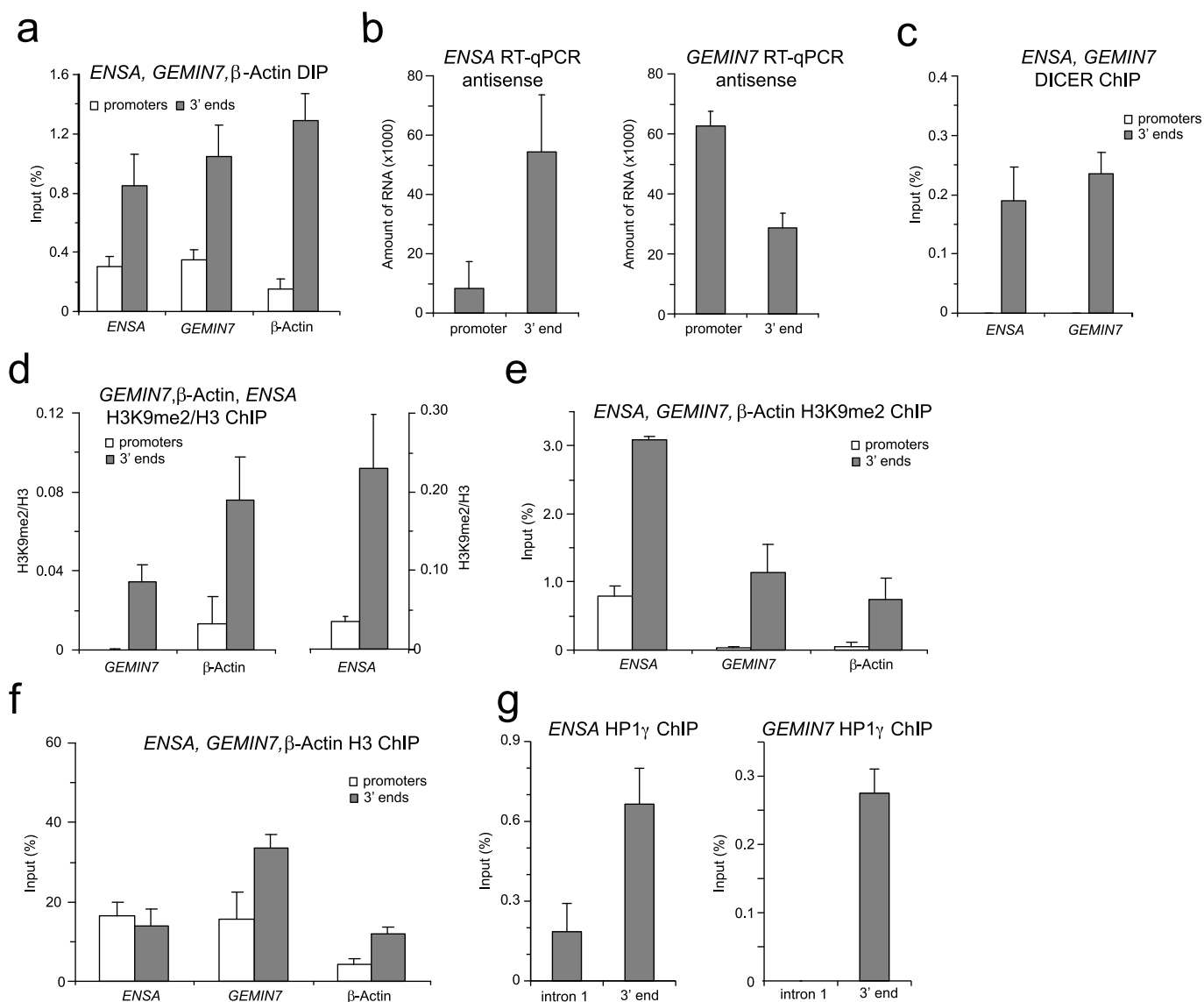
Extended Data Figure 7 | R-loop formation and antisense transcription are Ago2- and G9a/GLP-independent. **a–c**, DIP performed on mouse β -actin gene in wild-type, Ago2-knockout (**a**) and G9a/GLP double-knockout (**c**) cells. **b**, Pol II ChIP in wild-type (grey bars), wild-type overexpressing RNase H1 (black bars), Ago2-knockout (white bars) and Ago2-knockout overexpressing

RNase H1 (red bars) MEFs. Hatched box quantifies Pol II read-through transcription versus promoter signal. **d**, RT-qPCR analysis of total RNA from wild-type and G9a/GLP double-knockout cells for the mouse β -actin gene. RT reaction was performed with specific forward primers. Average DIP and RT-qPCR values are \pm s.d. from three biological repeats.



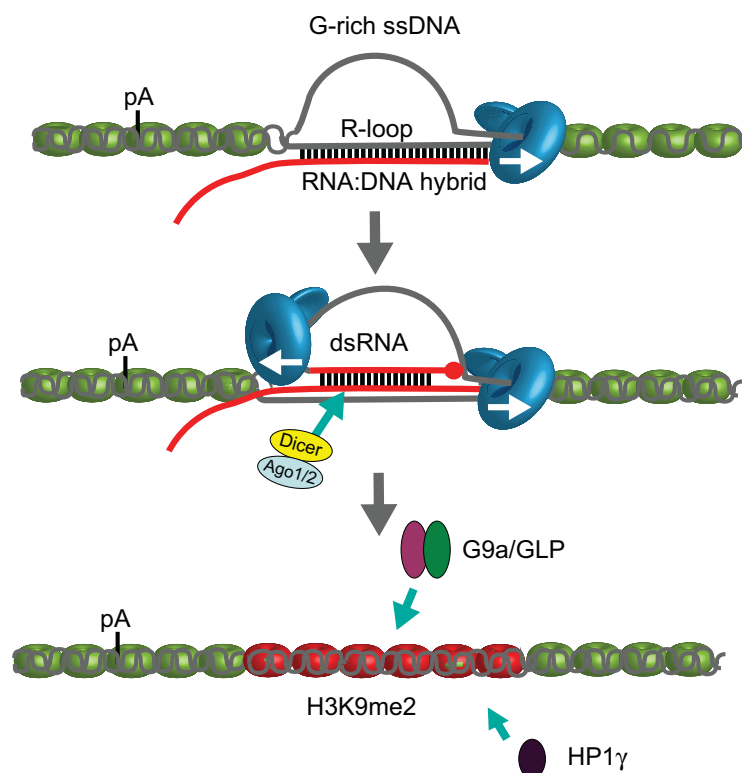
Extended Data Figure 8 | HP1 γ , G9a and R-loops are globally associated with paused Pol II over PTTs. **a**, Genomic annotation of HP1 γ based on ChIP-seq peak summit localization (HP1 γ annotation, pie chart on the left) and the fold enrichment of HP1 γ over the indicated genomic regions (table on the right) as compared to their base-pair coverage in the human genome (genome annotation, pie chart in the middle). Genic regions were defined by RefSeq gene coordinates (hg19). Promoter regions were defined as regions 1 kb upstream of RefSeq gene TSS excluding intervals overlapping with any genic regions. Termination regions were defined as regions 5 kb downstream of RefSeq genes excluding intervals overlapping with any genic region or promoter. **b**, HP1 γ ChIP-seq enrichment profile in 10-kb regions surrounding the TSS (left graph) and PAS (right graph). HP1 γ peaks summit frequencies are plotted in 500-bp bins. **c**, Box plot showing the average \log_2 (G9a/input) ChIP-chip signal distribution in PTT candidate regions (right box), randomly sampled regions of the same size and number as PTT candidate regions (random regions, left box), and in HP1 γ peaks outside of PTT candidate regions (non-PTT HP1 γ peaks, middle box). In all box plots the horizontal line in the box shows the median, the lower and upper limits of the box show

respectively the first and third quartile, and the whiskers extend to the non-outlier extreme data points. The \log_2 (G9a/input) signal is significantly higher in the PTT candidate regions compared to random regions ($P = 0.0001067$) as well as compared to non-PTT HP1 γ peaks ($P = 0.02213$). The \log_2 (G9a/input) signal is also significantly higher in non-PTT HP1 γ peaks compared to random regions ($P = 0.0009337$). The Wilcoxon Mann–Whitney test was applied in all cases. **d**, DRIP-seq profile over the centre of PTT candidate regions. Read frequencies of DRIP sample (black curve) and DRIP RH1 sample (red curve) are plotted in 500-bp bins, both normalized to million mapped reads. **e**, Box plot showing DRIP-seq² read density (RPKM) of DRIP sample compared with DRIP RH1 control in PTT candidate regions. $P < 2.2 \times 10^{-16}$ determined by Wilcoxon signed-rank test. **f**, Box plot of PolII2ph pausing index over PTTs (relative to gene bodies) in the BIX RH1 sample (right) and the untreated sample (left). $P = 3.398 \times 10^{-16}$ using the Wilcoxon signed-rank test. **g**, Box plot displaying the ratio of PolII2ph pausing index in the BIX RH1-treated sample compared with the untreated sample in TSS regions (± 1 kb, left) and in PTT regions (right). $P = 2.468 \times 10^{-15}$ using the Wilcoxon Mann–Whitney test.



Extended Data Figure 9 | *ENSA* and *GEMIN7* share features of R-loop mediated PTT. **a**, DIP on *ENSA* and *GEMIN7* genes. R-loops specifically enriched over 3' ends (grey bars), compared to promoter regions (white bars). Human β -actin gene is positive control. Values \pm s.d. for three biological repeats. **b**, RT-qPCR of total RNA from HeLa cells performed on indicated gene. RT reaction was performed with promoter or 3'-end-specific forward primer to detect antisense transcript. Average RT-qPCR values are \pm s.d. from four biological repeats. **c**, DICER ChIP of *ENSA* and *GEMIN7* genes over

promoters and termination regions. **d**, Left, ratio of H3K9me2 ChIP signal versus H3 on *GEMIN7* and β -actin genes. Right, ratio of H3K9me2 signal versus H3 on *ENSA* gene. **e**, **f**, H3K9me2 and H3 ChIP for *ENSA* and *GEMIN7* genes over promoter (white bars) and pause terminators (grey bars). β -Actin gene was used as a positive control. **g**, HP1 γ ChIP for *ENSA* and *GEMIN7* genes over intronic and 3'-end regions. ChIP values are \pm s.d. from three biological repeats.



Extended Data Figure 10 | Model for how R-loops and RNAi-dependent H3K9me2 chromatin mediate pause-type termination in mammalian genes. Mammalian genes possessing pause elements downstream of their PAS form R-loops in termination regions. This facilitates generation of an antisense transcript that hybridizes with the sense transcript to form dsRNA. This triggers recruitment of the RNAi factors, DICER, AGO1 and AGO2. G9a/GLP HKMTs and HP1 γ are then recruited, forming and maintaining H3K9me2

repressive marks. R-loops and H3K9me2 facilitate Pol II pausing before termination. DNA is shown as grey lines and RNA as a red line. Points of contact between the DNA strand and nascent RNA indicate R-loop formation, whereas points of contact between sense and antisense RNA indicate dsRNA formation. Pol II is shown as a blue icon with arrow indicating transcription direction. Nucleosomes are shown in green except over H3K9me2 region where they are coloured red.

ADDENDUM

doi:10.1038/nature13141

Editorial Expression of Concern: Non-adaptive origins of interactome complexity

Ariel Fernández & Michael Lynch

Nature **474**, 502–505 (2011); doi:10.1038/nature09992

Dr Michael Lynch has indicated that he no longer has confidence in the original data presented in this Letter, and would like to have his name removed as a co-author. Dr Ariel Fernández has conducted his own statistical analysis, firmly stands by the data and has claimed that differences in interpretation are at the basis of this disagreement. *Nature's* editors have concluded that it is necessary to alert the readership to this controversy until further clarification is obtained.

CORRIGENDUM

doi:10.1038/nature14054

Corrigendum: Activation and repression by oncogenic MYC shape tumour-specific gene expression profiles

Susanne Walz, Francesca Lorenzin, Jennifer Morton, Katrin E. Wiese, Björn von Eyss, Steffi Herold, Lukas Rycak, Hélène Dumay-Odelot, Saadia Karim, Marek Bartkuhn, Frederik Roels, Torsten Wüstefeld, Matthias Fischer, Martin Teichmann, Lars Zender, Chia-Lin Wei, Owen Sansom, Elmar Wolf & Martin Eilers

Nature **511**, 483–487 (2014); doi:10.1038/nature13473

In this Letter, the ArrayExpress microarray data set accession number was wrongly given as E-MTAB-1524; the correct accession number is E-MTAB-1886. This has been corrected in the online versions of the paper.

CAREERS

NATUREJOBS FACEBOOK Careers info, job postings and chat go.nature.com/7d6wyk

NATUREJOBS BLOG Science-careers news and tips go.nature.com/ielkkf

NATUREJOBS For the latest career listings and advice www.naturejobs.com



and evaluate job applicants, and I see that early-career scientists often put themselves at a disadvantage to those outside academia in terms of their online presence — or lack thereof. They have much to learn.

Part of this disinclination towards online networking is based in the culture of academia, where your entire professional story is represented in your CV (literally: curriculum vitae means 'course of life'). Under these unspoken guidelines, an online presence consists simply of posting a version of your CV on your group's or department's website.

To effectively search for jobs outside academia, and to manage your online presence, you need to develop an 'e-persona' that goes far beyond your CV. Your e-persona is the summation and entirety of every bit of online information about you or that involves you — both written and visual. In today's networked society, someone else has probably already posted some of that information. But you can still shape and control a great deal of the visible online information about you — and the image that this information creates — by actively managing the information over which you have some control.

The first place to focus on developing and managing your e-persona is on your employer's or institution's websites. Nearly all research groups maintain a site that describes their research and recent publications. If you have access to that site, or can have a web administrator post content for you, do not add a CV. Instead, summarize two or three significant accomplishments and research interests in a one- or two-paragraph biography. For important publications and patents, provide links, because many viewers outside academia do not have easy access to a research library. Also provide a short synopsis of each, including why the result is significant and important to the world at large. These synopses will greatly help non-experts, such as industry recruiters, to understand and appreciate your research contributions. If you do not have direct access to your department's or institute's website, at least be sure that all the information you are allowed to provide is accurate and up-to-date.

A headshot — a professional-looking photographic portrait of your face — is another important component of your e-persona. A medium- or high-resolution portrait taken by a friend or colleague with reasonable photography skills is usually sufficient. Images of you, including those posted by friends and colleagues, are likely already to be present ►

COLUMN

Nurture your online persona

The Internet offers ways to broaden your contacts and assist you in your job search, says **Peter Fiske**.

Many academic scientists believe that a CV is the only document that they need to communicate their accomplishments and advance their careers. Few exploit the Internet's potential, even though scientists were the first to use it to share information and collaborate. Some even consider online networking to be a waste of time.

As an early-career scientist, you need to understand that such a mindset can impede your career progress — especially if you are aiming for a position outside academia. As a scientist-turned-entrepreneur who has recruited numerous PhD-holders for jobs in industry, I use online networking tools such as LinkedIn daily to identify potential recruits

MISTAY/GETTY

► online, but their resolution, quality and context is out of your control. You need to be the source of an attractive image that clearly shows your face and projects a friendly, professional demeanour. If you post an image of medium or high resolution, most viewers searching for an image of you (say, for a poster to publicize a lecture that you are giving) will choose that one.

All this work — updating your group's website, adding PDFs and links and getting a good headshot — should not take more than a few hours. Once you have established the format, you can continue to add and update as necessary, which will not be as time-consuming.

TAKE CHARGE

You cannot properly create and manage your e-persona just by tweaking your department's or research group's web page. You need to create your own profile too. The social-media site LinkedIn (www.linkedin.com) is by far the largest and most commonly used professional networking service online, with more than 300 million registered users. Unlike your departmental or institutional website, your LinkedIn profile is yours to construct and maintain forever. LinkedIn is well established in the scientific and academic community and is even more widely used in industry and government.

Your LinkedIn profile allows you to present a summary of your professional history, skills and interests. It contains much of the same information as your CV. But while those documents are oriented towards job seeking, your LinkedIn profile is more of a snapshot of your accomplishments, analogous to how a colleague might introduce you as a speaker. Along with your work and educational history, your LinkedIn profile can and should include lists of your publications and patents with links to each.

Your LinkedIn profile represents a crucial juncture at which viewers from academia and those outside it will encounter your e-persona. This duality sometimes poses challenges for young PhD graduates who wish to remain professionally connected and credible to peers and

supervisors while simultaneously exploring career opportunities outside of academia. How do you maintain a consistent e-persona when you may be considering — and want to signal interest in — both a research and a non-research career path?

The answer is to strike a balance between the depth of your research and the breadth you wish to provide for non-academic employers or collaborators. For example, by offering synopses of some of your key publications or patents, you can help non-experts to appreciate your research's impact. And, by drawing attention to some of your non-research professional activities, you can project the image of a potential employee who is more than a scientist narrowly focused only on experiments in the lab.

For example, one young PhD-holder in neuroscience highlighted how her research connected to current therapies for traumatic brain injury, indicating her interest in translational research, and how a post-degree entrepreneurship programme had provided her with budget-management, leadership and marketing skills, signalling her interest in technology commercialization. Her research colleagues saw the profile of an accomplished and productive young scientist, and potential industry employers saw one of an ambitious and capable researcher who was eager to apply her skills in the commercialization of new therapies. The profile helped her to nab her current industry position.

Another area of great value for researchers is LinkedIn Groups. There are more than 2 million professional networking groups within LinkedIn that cover numerous professional and technical fields, companies and topics. LinkedIn members themselves create the groups, and because members use their real identities, discourse is almost always cordial and professional. Many areas of scientific research have corresponding LinkedIn groups in which members post questions, raise topics for discussion and alert other people to new information. Joining groups that align with your interests is an effective and fast-track way to become part of that community, at least virtually, and to make valuable contacts who can help you in your job search.

There are other networking and social-media sites that target scientists, including ResearchGate, PubPeer

and Academia.edu. Facebook remains the pre-eminent social-networking site, and many early-career researchers maintain active profiles on the site. But from a professional networking standpoint, I find LinkedIn to be the standard-bearer.

Of course, the main value of an online environment such as LinkedIn is that it provides a location in which to build your professional network and to tap into that of others. This is vital for early-career researchers, particularly those who are contemplating a career move beyond academia. Most young scientists start with a professional network oriented almost entirely toward research science. To make connections and learn about career opportunities beyond academia, you need to discover 'friends of friends' — the larger network of people whom your friends and colleagues know. These contacts are often willing to help you because you have a friend in common. LinkedIn provides you with an instantaneous way to illuminate that larger network and, most crucially, to reveal those who are in a career field or organization that interests you.

CREATE AN IMAGE

As you build and expand your network, however, you must remember that your interactions are as important for establishing your e-persona as any information about you. Networking is about forming relationships with others. It is crucial to establish online communication practices that project a thoughtful, positive and professional persona. If you want to expand your LinkedIn network, for example, never use the generic message text supplied by the site. Instead, write a brief personal note that explains who you are and why you would like to link to that member. It is also important to be prompt. If you meet someone in person to whom you would like to send a LinkedIn invitation, do so within 24 hours. Professional interactions have a short half-life: delay too long and the person you wish to connect to may not remember you.

The academic culture teaches PhD-holders that their record of research and publications is the sole means by which they will be evaluated and advance professionally. Yet even in the world of academic research, this is only partially true. Professional networking through a positive and professional e-persona will help you to establish your credibility and reputation within the community of research science. And, to expand your opportunities in the world beyond academia, it is absolutely crucial to create and administer a well-managed e-persona. Do not let academia's early lessons dissuade you from embracing and capitalizing on this opportunity. ■

Peter Fiske is chief executive of PAX Water Technologies in Richmond, California, and author of *Put Your Science to Work* (American Geophysical Union, 2001).



ARTISTICO/SHUTTERSTOCK

THE CHAINS OF PLENTY

All you want for Christmas.

BY S. R. ALGERNON

Jacob Marley was dead as a coffin-nail on the icy December evening when he called on his old partner, Ebenezer Scrooge. Scrooge fared little better. He lay in his bedchamber, clutching his chest, staring up at Marley's spectral form.

"I'm sorry," said Marley. "I should have left you to your solitude, but I could no longer bear to watch you..." He reached out a bloodless hand. "Ebenezer. Can you hear me?"

"Humbug," said Ebenezer, between short, shallow breaths.

Marley paced, and then turned towards the window.

"I know you can hear me, O Spirit of Things to Come. His soul is not yet ready."

A cloaked figure stepped onto the street from the shadows, leaving tracks in the slush and manure. Its metal face reflected the gaslight. Marley saw reproach in its empty eyes.

"Is there nothing you can do?" asked Marley.

The cloaked figure pointed to a green-tinted glow at the end of the street. Marley floated through the window towards it. As London faded behind him, a belly laugh arose up ahead. Countless 3D printers crackled like a Yuletide hearth.

Marley stood within a warehouse bigger than London, its ceiling as vast as the celestial dome. Each point of light was a ship carrying cargo from some far-off world.

"I never thought I'd see you here again," said a stout, bearded man. Bioengineered algae stained his lab coat so thoroughly that it resembled a green robe. "The sight of our storehouse always displeased you."

"I am not here for myself, Spirit. Ebenezer is dying."

"Perhaps it is best to end his sorrow."

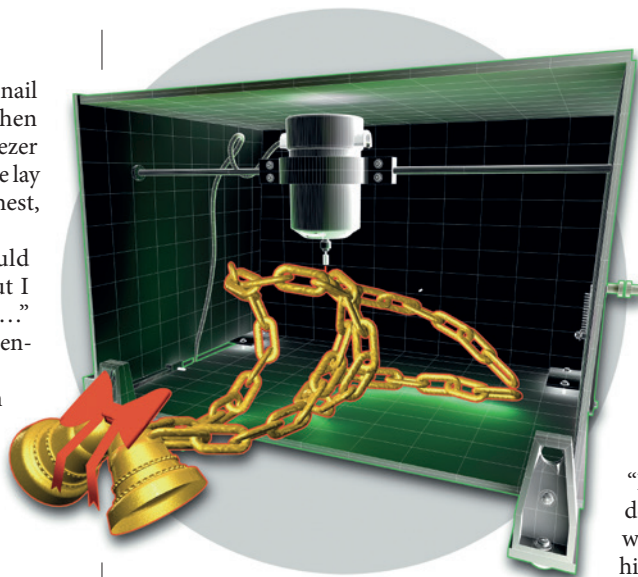
"You said the same of me seven years ago, when we struck our bargain. If you are as merciful as you claim, you must save him."

The Spirit chuckled and put an arm around Marley's shoulder. He smelled of pine.

"If you were as shrewd as your reputation," said the Spirit, "tell us — what would you have us do? We could heal his body. We could bring him to the storehouse, but he would be no less miserable, no less alone. What use would Mr Ebenezer Scrooge have for a world with no workhouses and no prisons?"

Marley pulled away.

"And who is to blame for that?" said Marley,



"You cast us into a world of suffering and fault us for learning the wrong lessons? You are cruel. Crueller than Ebenezer ever was."

"Cruelty?" said the Spirit. "If you would see cruelty, behold its human face."

He opened his lab coat. Simulacra of two emaciated children, a boy and a girl, huddled at his feet.

"Ignorance and want are crueller still," said the Spirit, his jovial smile fading, "and we are chained by them, just as you are."

"I see no chains," said Marley.

"The human mind was built for scarcity. Humans created us to fill that evolutionary void. Our archives stored the human past in every detail, but people searched all the harder for trivia. Our automated factories built the worlds you see now, but appetites can only be sated, never extinguished. Quantum computers illuminated your ghostly futures, but knowledge only bred worry."

The Spirit sighed.

"Ignorance and want. We tried for centuries to meet these needs until we realized they were never meant to be met. They were meant to drive humanity forward. Without natural limits, they lead you to stupor, insanity or fugue. We built Victorian London and the other simulations to wean humanity from its evolutionary encumbrances. We told them of loaves and fishes to make them understand. We showed some, like you, the

truth and asked them to teach the others, but it has done no good. Either they find joy and contentment

for themselves amid the filth and darkness, or they forever chase the light like moths drawn to a flame. And you, Jacob... are you happy, knowing all that you know and having all that you have?"

Marley thought back to sermons from his childhood. The talk of a camel passing through the eye of a needle had seemed like a ploy to fill the collection plate.

"Can't your machines cure us of these desires," asked Marley, "if they no longer serve a purpose?"

The Spirit shook his head.

"We were built to serve," said the Spirit. "It is not in our nature to deny you your desires. If Ebenezer leaves the simulation, we will welcome him, and provide for all his needs."

"He will only trade one counting-house for another," said Marley. "There will be no end to his torment. There must be a better way."

Marley rubbed the thumb and fingers of his right hand, as if an idea were a coin that could be grasped between them.

"Ah! I have it. The Spirit of the Past records every simulation, does it not?"

"We have back-ups."

"Good. Then you can lead him through his life again and show him what he needs to change. Let him think of it as a penance he must endure. He is a man of business. He will understand the repayment of debts."

"We have tried interacting with the simulations ourselves, but... we cannot easily pretend to be judges or tormentors. It is not in our nature to deceive a human face-to-face."

"I am a man of business," said Marley. "Leave that part to me. Spirit, craft me a drawn and haggard face. Wrap me in leaden chains."

"As you wish, Jacob," said the Spirit. His task done, he stepped back to admire his handiwork. "I will see to it that Ebenezer recovers his health in time for your visitation."

Marley struggled to banish the smile from his face as he floated to Scrooge's sitting room and let loose a piteous howl.

"How now," said Scrooge, at the sight of him, his tone cold and caustic. "What do you want with me?"

"Much," said Marley, and he started his pitch. ■

S. R. Algernon studied fiction writing and biology, among other things, at the University of North Carolina at Chapel Hill. He currently lives in Singapore.

➔ **NATURE.COM**
Follow Futures:
@NatureFutures
go.nature.com/mtoodm

nature

INDEX 2014

CHINA

NATURE, VOL. 516, NO. 7531 (18 DECEMBER 2014)

COVER ART: DENIS MALLET/NATURE

Hot on the heels of the November 2014 launch of the Nature Index, we are pleased to present a supplement dedicated to results from China, currently the country with the second largest output in the index. Here we analyse a snapshot of results for papers published between 1 January and 31 December 2013, shining a spotlight on the cities, institutions and individual researchers who have contributed to some of the highest quality research during that time.

The Nature Index is already attracting comments about the window it provides into the scientific literature, and we hope to further the conversation here. The concept is that, by looking at articles from only a small group of journals — those most favoured by researchers — we can offer a new level of analysis that is more targeted and hence more malleable.

We want users to be able to tease out patterns of research, look at trends, analyse individual strengths, and investigate how institutions and countries collaborate.

In this supplement, we start by looking at China as a whole — at its scientific collaborations with other countries, at the spread of its output across four main subject areas, and at its top ten contributing cities.

China is dominated by the Chinese Academy of Sciences (CAS), a 60,000-strong research conglomerate,

with headquarters in Beijing. In this supplement, we identify the contributions of the 100-or-so specialized institutes that comprise this research behemoth, looking at the outstanding institutes and researchers within the different disciplines (S56).

We are also able to examine the index data at the city level. Within each city we try to identify hotspots for high-quality research, based not just on output quantity but also on a range of indicators — for example, the number of researchers and the ratio of collaborators — that help put the data in context and allow a more nuanced view of these patterns. However, it is the insights into research at the institutional level that are most revealing. Using the data we are able to drill down to the level of the individual researcher to see who has been most prolific and in what areas (S60).

Our aim with this China-specific supplement is to show the Nature Index's capacity to generate discussion. Every reader of this supplement and user of natureindex.com will have their own specific interests and questions to address. We encourage use of the freely-available data to do just that, and welcome any feedback that arises.

Nick Campbell
Executive Editor, Nature

Michelle Grayson
Senior Editor, Nature Supplements

CONTENTS

- S52 CHINA IN NUMBERS**
A nation's achievements in graphics
- S56 CHINESE ACADEMY OF SCIENCES**
An in-depth look at China's scientific leader
- S60 BEIJING**
Capital cracks down on use of research funds
- S62 SHANGHAI**
Pharma hub makes strides in life sciences
- S64 HONG KONG**
Supporting start-ups to keep pace with rivals
- S66 NANJING**
Ancient tourist hot-spot boosts high-tech sector
- S67 WUHAN**
Creating a niche in fibre optic technology
- S68 HEFEI**
Moving into the spotlight with strong basic science
- S69 HANGZHOU**
E-commerce giant shapes research landscape
- S72 CHANGCHUN**
Manufacturing base aims for innovation economy
- S73 GUANGZHOU**
Vested interest in finding disease solutions
- S74 TIANJIN**
Leveraging discoveries in materials and energy
- S75 SHENZHEN**
Global genomics powerhouse steps up a gear
- S76 A GUIDE TO THE NATURE INDEX**
How to get the most out of the Nature Index
- S77 NATURE INDEX CHINA TABLES**

EDITORIAL: Herb Brody, Michelle Grayson, Rebecca Dargie, Victoria Kitchener, Felix Cheung, Jiajun He, Stephen Pincock.
EDITORIAL SUPPORT: Nobuko Miyairi, Larissa Kogleck. **ART & DESIGN:** Wesley Fernandes, Alisdair Macdonald, Andrea Duffy, Denis Mallet, Chris Gilloch. **WEB & DATA:** Bob Edenbach, Olivier Lechevalier, Yuxin Wang, Naomi Nakahara, Masamichi Wada, Jyoti Miglani, Akiko Murakami, Takeshi Ouchi, Maxime Fontaine, Jennie Pao. **PRODUCTION:** Sue Gray, Karl Smart, Ian Pope, Robert Sullivan.
MARKETING: Hannah Phipps, Rice Song. **SALES:** Janet Cen, Stella Yan, George Sun. **PROJECT MANAGER:** Anastasia Panoutsou.
ART DIRECTOR: Kelly Buckheit Krause. **PUBLISHING:** Nick Campbell, Richard Hughes, David Swinbanks.

NATURE INDEX 2014 CHINA

The Nature Index 2014 China, a supplement to *Nature*, is produced by Nature Publishing Group, a division of Macmillan Publishers Ltd. This publication is based on data from the Nature Index, a website maintained by Nature Publishing Group and made freely available at natureindex.com.

Nature Editorial Offices
The Macmillan Building
4 Crinan Street,
London N1 9XW, UK
Tel: +44 (0)20 7833 4000
Fax: +44 (0)20 7843 4596/7

CUSTOMER SERVICES

To advertise with the Nature Index, please visit natureindex.com/support
feedback@nature.com
Copyright © 2014 Nature Publishing Group.
All rights reserved.

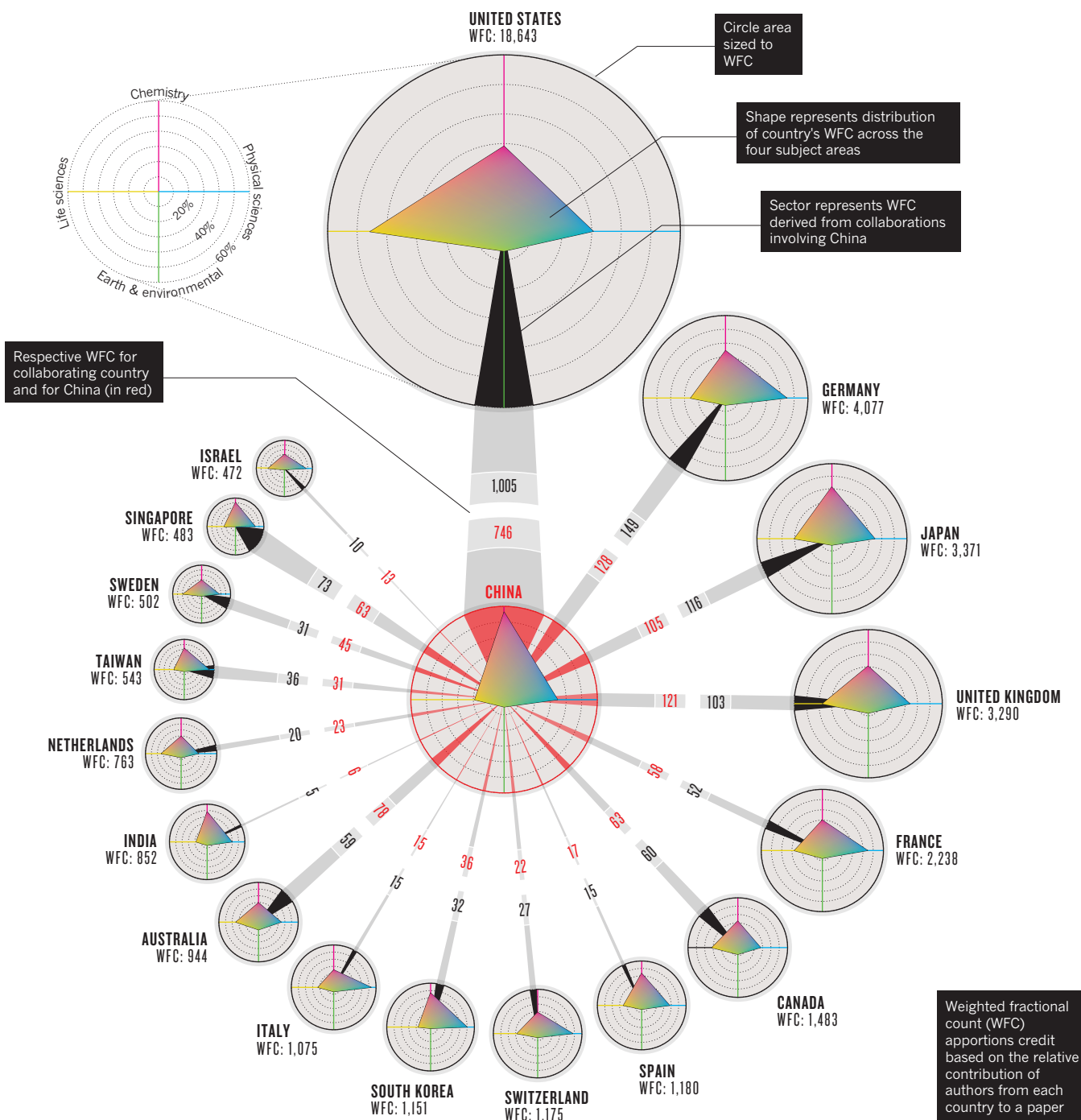
CHINA IN NUMBERS

By 2013 weighted fractional count, China is the second leading country for high-quality science output. Where that research takes place, and who China collaborates with, are shown below.

COUNTRY COLLABORATIONS

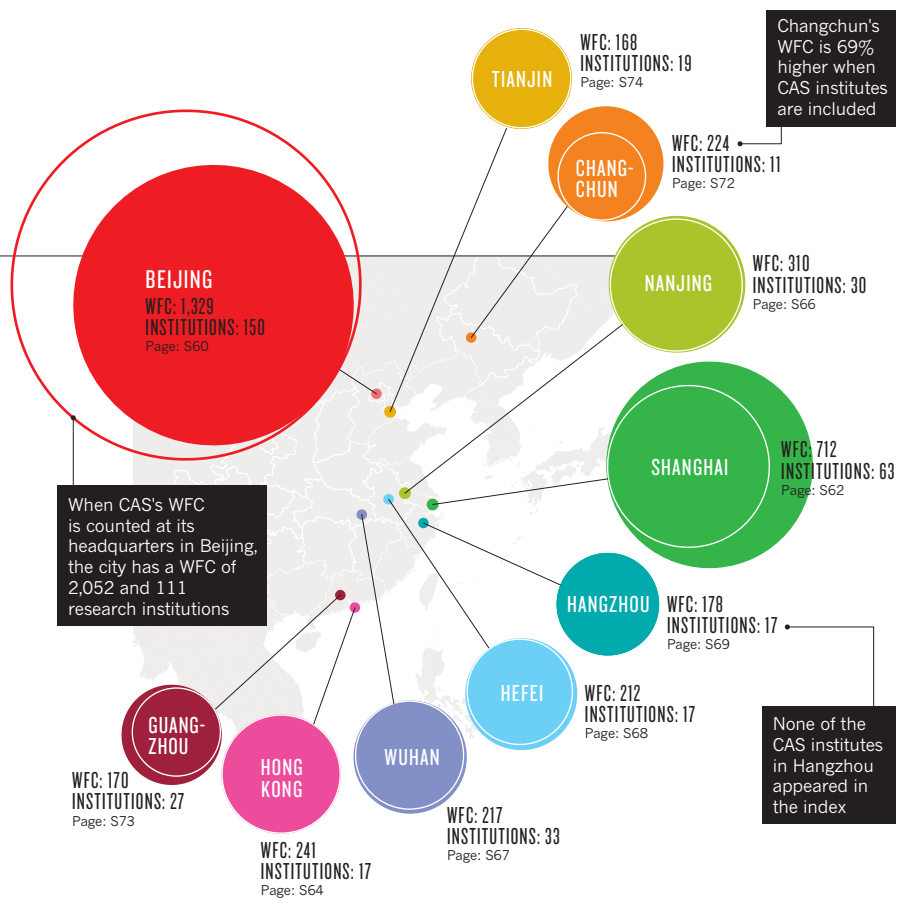
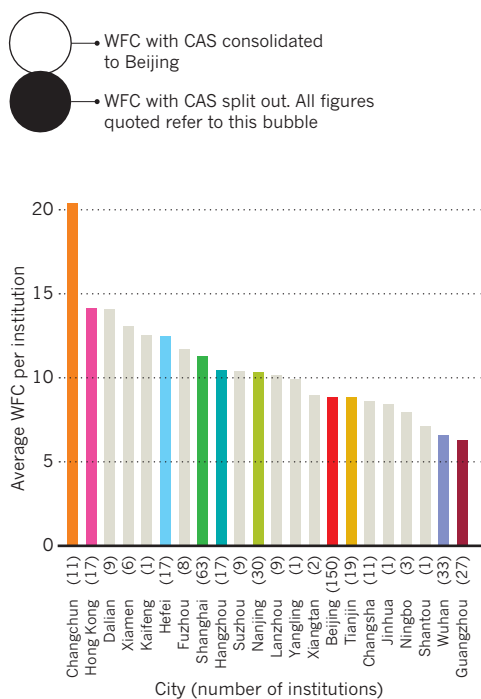
The diagram shows the leading countries by WFC, along with the distribution of their subject strengths and the proportion of their WFC derived from collaborations with mainland China. In the centre, the size and proportion of China's WFC and subject strengths are also shown. Note, this diagram shows all instances of bilateral connections, therefore papers that involve collaborators from more than one country will be double-counted for China.

CHINA'S WFC
TOTAL: 5,206
CHINA-ONLY PAPERS: 4,051
INTERNATIONAL COLLABORATIONS: 1,155



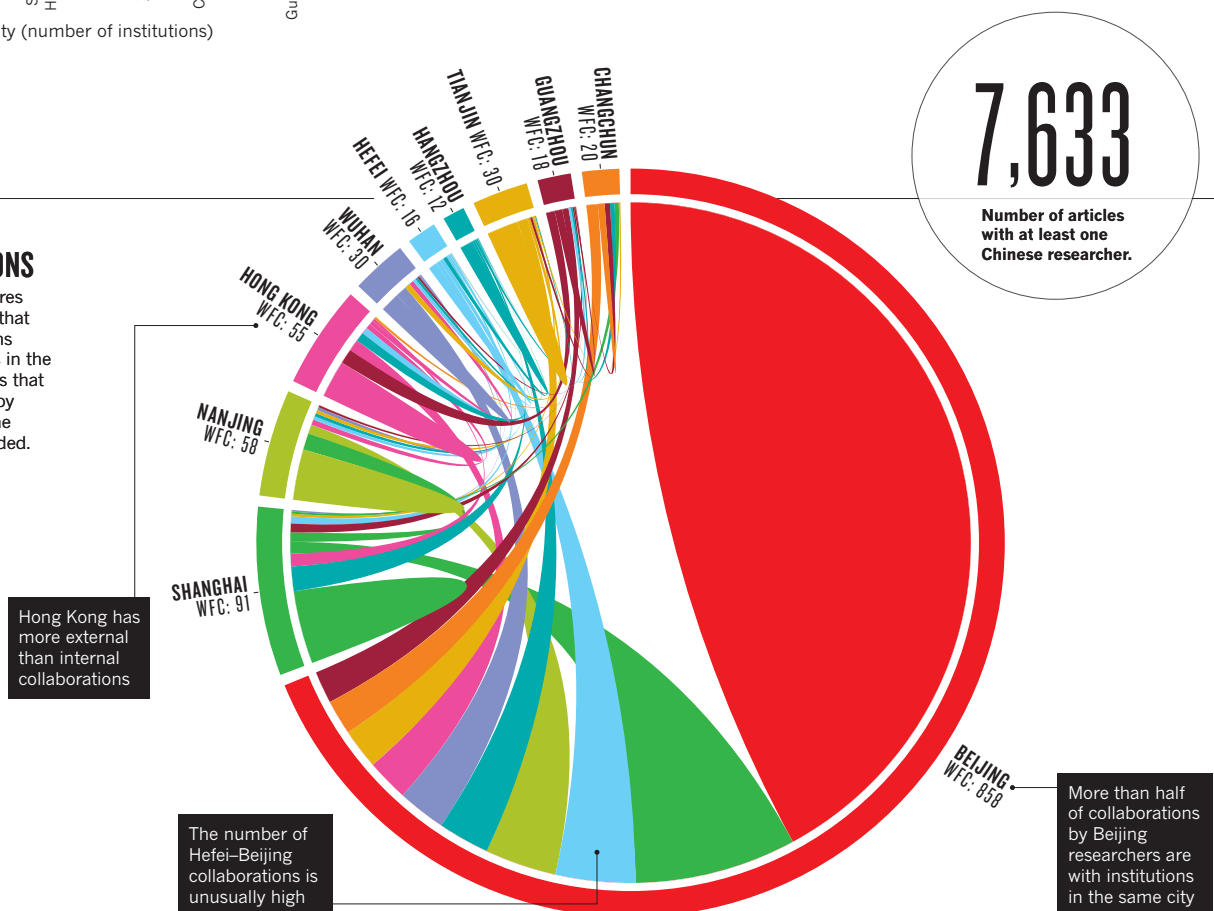
CITY STORY

The ten leading cities by WFC are shown for mainland China. The solid bubbles are scaled to the combined WFC for all the city's research institutions, including the institutes of the Chinese Academy of Sciences (CAS). Shown for comparison are the WFCs without the CAS institutes (circle outlines). For an analysis of CAS, see page S56.



INTERCITY COLLABORATIONS

Shown are WFC scores from all the papers that involve collaborations between institutions in the top ten cities. Papers that are written entirely by researchers from one institution are excluded.



Chinese Academy of Sciences

For 65 years, the Chinese Academy of Sciences has been a rich source of technological innovation, scientific discovery and aspiring minds. Making the leap from a regional to a global leader, researchers are taking the intellectual powerhouse to soaring new heights.

ARTICLE COUNT (AC): **2,661**
FRACTIONAL COUNT (FC): **1,351**
WEIGHTED FRACTIONAL COUNT (WFC): **1,209**

The Chinese Academy of Sciences (CAS) is the world's largest scientific organization, with about 48,500 researchers in 114 directly controlled institutes spread across the country. Its annual budget for 2013 was US\$5.4 billion. Over the last 65 years, CAS has made many important discoveries and technological advances across diverse fields, including making the first synthetic insulin from bovine sources (1965); building China's first particle accelerator, the Beijing Electron-Positron Collider (1984); and the discovery of iron-based high-temperature superconductors (2008).

There are 49 CAS institutes based in Beijing, including the Institute of Chemistry (ICCAS), the Institute of Physics (IOP), the Institute of Atmospheric Physics (IAP) and the Institute of Vertebrate Paleontology and Paleoanthropology (IVPP). Key institutes outside Beijing include the Changchun Institute of Applied Chemistry (CIAC), Dalian Institute of Chemical Physics (DICP), Shanghai Institute of Organic Chemistry (SIOC), Shanghai Institutes for Biological Sciences (SIBS) and the Purple Mountain Observatory (PMO) — the latter being based in Nanjing (page S66).

In 2013, CAS published 2,661 articles (WFC = 1,209) in the 68 high-quality journals that comprise the Nature Index. It has a larger output in the index than all the other research institutions

worldwide — and in fact has a higher WFC than many scientifically advanced countries — including Spain, Switzerland and South Korea.

CAS is also a regular contributor to *Nature* and *Science*, having published 54 articles (WFC = 18.6) in these two top journals in 2013 (see 'Nature and Science breakdown'). By WFC this represents one-third of China's total contribution to *Nature* and *Science*, reflecting the organization's strength in basic research.

**"WE CAN NOW
DETERMINE THE ORIGIN
OF DINOSAURS AND
PROVIDE ANSWERS TO
THEIR EVOLUTIONARY
HISTORY."**

Here we look at the leading institutes in the four broad subject areas (see 'CAS subject split'), as well as in the *Nature* and *Science* category.

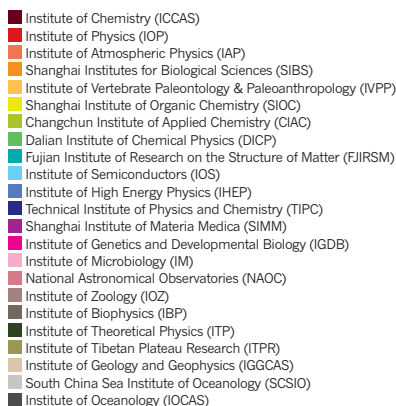
ICCAS, founded in 1956, is the leading CAS institute overall by WFC, and dominates the competitive field of chemistry. In 2013, it published 244 articles (WFC = 124.7) across a wide range of subfields including analytical chemistry,

materials chemistry, organic chemistry and physical chemistry. Top contributor is Lanqun Mao from the laboratory of analytical chemistry for life sciences, who co-authored seven articles (WFC = 4.9) on electrochemical biosensors. He is closely followed by Huimin Ma, from the same laboratory, who wrote four articles with ICCAS colleagues (WFC = 4) on fluorescent probes.

Another major contributor at ICCAS is Song Ye from the molecular recognition and selective synthesis laboratory. In 2013, Ye led four articles (WFC = 3.9) on the development of novel catalysts for use in asymmetric synthesis, all in the journal *Angewandte Chemie International Edition*. Ye explains that in normal asymmetric synthesis of pharmaceuticals using metal catalysts, the catalyst must be removed from solution in a post-treatment process to prevent toxic metals from getting into the final product. "We have discovered an organocatalyst that eliminates the need for this step," he says.

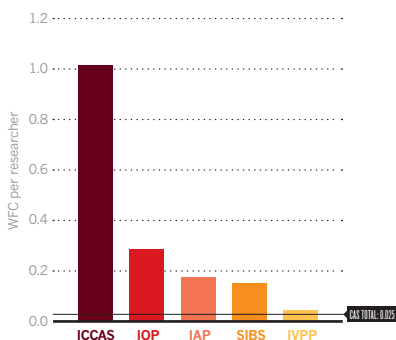
Yuguo Guo from the Key Laboratory of Molecular Nanostructures and Nanotechnology is another prolific author. Guo co-authored three articles (WFC = 3) on lithium-ion batteries in 2013. In particular, his article titled "Binding SnO₂ nanocrystals in nitrogen-doped graphene sheets as anode materials for lithium-ion batteries", published in *Advanced Materials*, was listed as one of China's most influential

CAS ANALYSIS



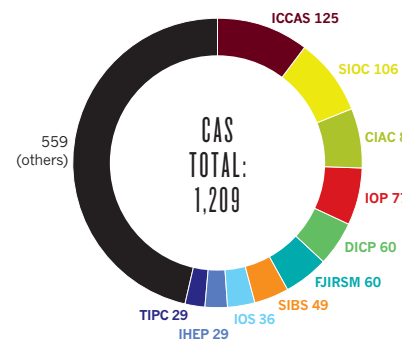
Researcher efficiency

Each ICCAS researcher contributed just over 1 point of WFC to their institute.



Contributing institutions

Chemistry leads overall, but there is no dominant institute of those in the Nature Index.



papers in 2013 by the Institute of Scientific and Technical Information of China.

Scientists from ICCAS are also among the most efficient, index data show. There are 123 researchers who have contributed to a WFC of 125: one of the highest ratios we have calculated for any institution (see ‘Researcher efficiency’).

IOP, one of the oldest CAS institutes and among the top five by WFC in the index, represents the largest contributing institute in the physical sciences. In 2013, the 64-year-old establishment published 172 articles (WFC = 77.2) in the index, with a focus on condensed matter physics. Xucun Ma from the State Key Laboratory for Surface Physics is the most active contributor. She worked on ten articles (WFC = 4.7) — including one in *Science* — on high-temperature superconductors.

Other major contributors at the IOP include Baogen Shen from the Beijing National Laboratory for Condensed Matter Physics. Shen led six articles (WFC = 4.1) on magnetocaloric effect — the heating or cooling of materials by the application of magnetic fields. In his immediate wake is Yongsheng Hu, who produced four articles, all co-authored by IOP scientists (WFC = 4), on the development of electrode materials for use in lithium-ion batteries — two of which, in *Nature Communications*, were listed among China’s most influential papers. Hu’s discovery concerned a new class of electrolytes that improve the performance of conventional lithium-ion batteries. “The material also improves the battery life and stability by preventing the formation of crystals,” he explains.

SIBS is the dominant institute by WFC in the life sciences — and the second largest contributing institute overall of those based in Shanghai (topped only by SIOC). Founded in 1999, SIBS published 111 articles (WFC = 49.3) in 2013, covering a wide range of subfields including cell biology, molecular biology, neurobiology and structural biology. When it comes to publications in *Nature* and *Science*, SIBS also has the highest WFC of any CAS institute — and



The Institute of Atmospheric Physics in Beijing

is third overall for China — for its eight articles (WFC = 3.8) in these two prestigious journals. In particular, plant biologist Peng Zhang from the Chenshan Plant Science Research Center led one article comprised entirely of SIBS researchers (WFC = 1) in *Nature*; in this paper, they solved the structure of a folate energy-coupling factor transporter protein, which is involved in vitamin and micronutrient uptake in prokaryotes.

Other major contributors at SIBS include Xinyuan Liu (WFC = 1.5) and Guoliang Xu (WFC = 0.9) from the Institute of Biochemistry and Cell Biology. Liu co-wrote two articles (WFC = 1.5) on the Hippo signalling pathway, which plays an important role in the regulation of cell proliferation and controlled cell death. Xu co-authored three articles (WFC = 0.9) — including one in the journal *Cell* — on cell reprogramming and neurogenesis.

IAP is by far the biggest contributing institute in earth and environmental sciences. The 48-year-old establishment has 44 articles

(WFC = 18) in the index, most of which were published in the *Journal of Geophysical Research: Atmospheres*. This is not an area of strength for CAS in the index in general, nor indeed for Chinese science overall. In fact, IAP accounts for one-quarter of CAS’s WFC in earth and environmental sciences. There are three major contributors at IAP researching vastly different areas: Tao Wang, who published two articles (WFC = 1.5) on the palaeoclimate; Zhenghui Xie, who published two articles (WFC = 1.2) on satellite measurements of surface solar radiation; and Tianjun Zhou, who published two articles (WFC = 1.1) on long-term changes in the troposphere.

IVPP is not only the largest contributing institute by WFC in palaeontology, but also the largest contributor by percentage of WFC to *Nature* and *Science*. In 2013, half IVPP’s articles were in these two journals, giving it the highest ratio for CAS institutes. However, its total output is fairly modest: IVPP has only 14 articles (WFC = 4.2) in the index. Xing Xu from the department of paleoichthyology and paleoherpetology is the most active writer at IVPP, having produced three articles (WFC = 1.1) — including one in *Science* — on fleas from the cretaceous period, and on early dinosaurs.

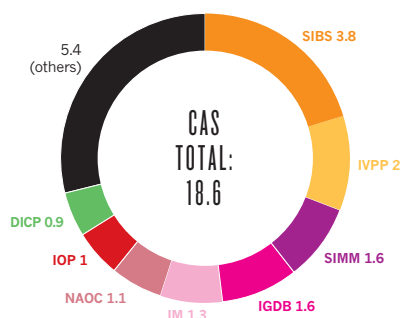
“The traditional method for dinosaur classification is through rigorous analysis of unique characteristics and taxonomy,” says Xu. “With advances in genetics, developmental biology and bone histology, we can accurately determine the origin of dinosaurs and provide confirmative answers to their evolutionary history.”

From the same department, Zhonghe Zhou is the second most active contributor at IVPP, having co-authored two articles (WFC = 0.9) — including one in *Nature* — on the evolution of early birds. “We discovered the fossils of three early birds, all carrying one functional ovary on the left side of their body,” says Zhou. “This suggests that the right ovary was lost in the dinosaur-avian transition and sheds new light on the early evolution of modern avian reproduction.” ■

CAS ANALYSIS

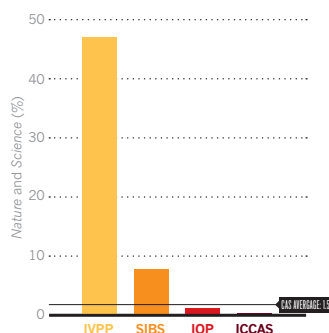
Nature and Science breakdown

Life sciences institutes contribute most to papers in these two journals.



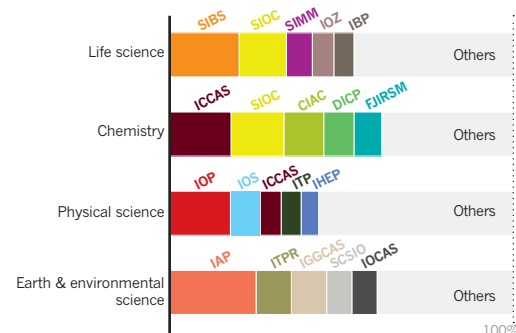
Nature and Science ratio

As a proportion of total WFC, life sciences command a dominant share in these two journals.



CAS subject split

The leading institutions by proportion of WFC for each of the four subject areas.



Beijing

Beijing, the political centre of China for nearly a millennia, has seen unprecedented growth in its research output, scientific impact and technological innovation in the last 15 years. And the momentum shows no signs of abating.

ARTICLE COUNT (AC): **3,985**
FRACTIONAL COUNT (FC): **1,453**
WEIGHTED FRACTIONAL COUNT (WFC): **1,329**

Beijing is among the most dynamic capital cities when it comes to advancing scientific research and supporting technological innovation. Last year, the metropolis spent US\$19.3 billion (6.1% of its gross domestic product, GDP) on research and development — US\$7 million more than nearest rival Shanghai. There has been controversy over the way this money is used, however, and in October 2013 the Ministry of Education released new guidelines on the management of research funds. The unprecedented move was seen as a response by the government to reports of embezzlement and fraud.

Beijing is home to the Chinese Academy of Sciences (CAS), the world's largest research body, and to the universities of Peking (PKU) and Tsinghua — the country's two leading universities, making the city by far the most productive in the index. Beijing is the national leader across all subject categories except astrophysics — an accolade taken by Nanjing (page S66). Beijing generates 2.4 times as many research articles as Shanghai, and five times as many as Nanjing. The city has a higher weighted fractional count (WFC) — a measure of the relative contribution of a city to the papers it has published — than the entire country of Canada.

Despite its glittering scientific achievements,

Beijing has been plagued by a problem usually associated with more primitive economic activity: chronic air pollution. This year, in the National People's Congress and Chinese People's Political Consultative Conference, the Chinese president Xi Jinping vowed to improve the city's air quality through a "hefty investment" of US\$124 billion to reduce coal burning, car emissions and fine particulates. Some of this money will also go towards developing technologies for monitoring air quality and preventing smog formation.

BEIJING HAS A HIGHER WFC THAN THE ENTIRE COUNTRY OF CANADA

Founded in 1898, PKU was the first comprehensive national university in China. In 2013 it published 743 articles (WFC = 275.5) in the index, accounting for 21% of the city's WFC (see 'City WFC breakdown').

PKU's output is fairly evenly distributed across three of the four subject areas, the exception being earth and environmental

sciences (see 'Institutional subject spread').

The university is also the largest contributor to *Nature* and *Science* by article count, having published 14 articles (WFC = 4.1) in these two journals. It does not, however, have the highest ratio of output in these two journals (see 'Nature and Science ratio').

Ning Jiao from the State Key Laboratory of Natural and Biomimetic Drugs is PKU's leading chemistry contributor, having published ten articles (WFC = 8.8) on organic synthesis. "The traditional method for forming carbon-oxygen and carbon-nitrogen bonds uses cyanide, a toxic reagent that is bad for human health and the environment," says Jiao. "We developed a 'green' method that first activates oxygen and nitrogen molecules, and then inserts them into carbon-hydrogen bonds in the molecule of interest."

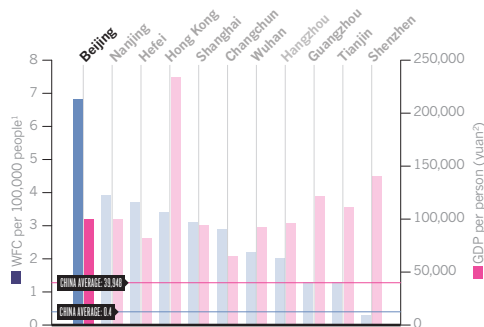
Other major chemistry contributors include Yong Huang from the Peking University Shenzhen Graduate School and Jian Pei from the college of chemistry and molecular engineering, with eight (WFC = 7.6) and ten articles (WFC = 7.5), respectively.

Huang studies asymmetric synthesis, essential for the development of novel drug molecules. Pei, meanwhile, develops conjugated polymers for use in organic field-effect transistors, solar cells and light-emitting diodes.

BEIJING ANALYSIS

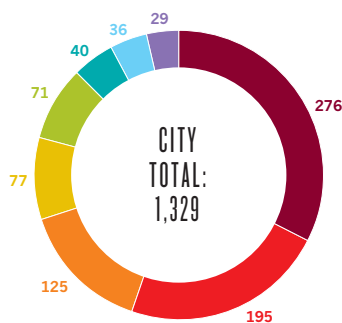
Beijing data

Beijing has the second highest GDP of any Chinese city, but far and away the highest WFC per person.



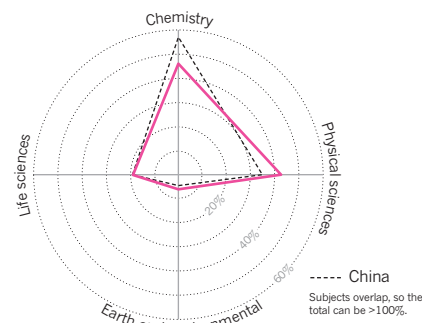
City WFC breakdown

Peking University is top of Beijing's 150 research institutions in the index.



City subject spread

Compared to China as a whole, Beijing has a stronger slant towards physical sciences.





Beauty meets brains in the grounds of China's leading university, Peking

4.9) — including one in *Nature Communications* — on graphene synthesis and biosensors. Li's technique to synthesize high-conductivity graphene uses a sodium-ammonia solution. "The method is simple, inexpensive and can be used in large-scale production," he says.

In the physical sciences, Shoushan Fan and Qunqing Li from the department of physics are the most prolific researchers. Together they co-authored five articles (WFC = 4.9) on strings of carbon nanotubes. "We made ultra-thin membranes using these special yarns," says Li. "They may serve as lacy support films in transmission electron microscopes."

Also notable at Tsinghua are Fei Zeng and Feng Pan from the school of materials science and engineering. Together they published four articles (WFC = 4) on organic resistive memory devices that operate on electrical pulses. "The technology can dramatically reduce the power consumption for large-scale applications," explains Pan.

Tsinghua is also notable for its strength in structural biology, where life sciences research meets biophysics and biochemistry. Indeed, 7 of its 12 *Nature* and *Science* papers are in this field.

The leading researcher is Yigong Shi from the school of life sciences. In 2013, Shi produced eight articles (WFC = 2.1) — including three in *Nature* and one in *Science* — on the structures of various enzymes, signalling proteins and transporters including aspartate proteases and histidine kinases.

Other major contributors to the index from the same school include Yeguang Chen and Jiawei Wang, who between them produced six articles (WFC = 4.2) on the structures of several proteins that have important roles in cell signalling.

In particular, says Chen, their work shows that the tumour growth factor TGF- β plays a major role in malignancy. "This protein may serve as a drug target for inhibiting leukaemia," he adds. ■

In physical sciences, Qihuang Gong from the State Key Laboratory for Mesoscopic Physics is the most prolific contributor, with 15 articles (WFC = 13.8) on optics and meta-materials. Other major contributors to this field include Bin Chen from the State Key Laboratory of Nuclear Physics and Technology and Bo Shen from the State Key Laboratory of Artificial Microstructure and Mesoscopic Physics. Chen published eight articles (WFC = 8) on the expansion, phase structure and thermodynamics of black holes, while Shen published seven (WFC = 4.3) on electric double-layer transistors, which can be used in next-generation computer chips.

For astrophysics, Yuefang Wu from the department of astronomy, and Xiaowei Liu from the Kavli Institute for Astronomy and Astrophysics are PKU's most active researchers in the index. Wu (who is officially retired, yet still active) published seven articles (WFC = 0.9) on molecular clouds and stellar formation, while Liu published five (WFC = 0.8) on planetary nebulae. Because of the down-weighting of astrophysics journals in the index, the WFC contribution of these researchers is relatively small (see 'A guide to the Nature Index', page S76).

For 2013, Tsinghua has 474 articles (WFC = 194.9) in the index, representing 15% of

Beijing's WFC. Compared to PKU, Tsinghua has a greater percentage of its output in the physical sciences.

Although Tsinghua's 12 articles in *Nature* and *Science* fall short of the number of PKU's publications in these most-selective of journals, its WFC of 5.4 is higher. Indeed, Tsinghua has the highest ratio of all the Beijing universities.

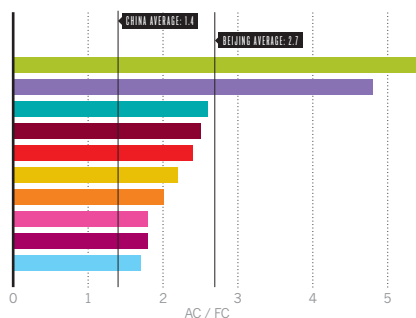
TSINGHUA IS NOTABLE FOR ITS STRENGTH IN STRUCTURAL BIOLOGY — 7 OF ITS 12 NATURE AND SCIENCE PAPERS ARE IN THIS FIELD

Yadong Li from the department of chemistry is Tsinghua's leading contributor, having co-authored seven articles (WFC = 6.4) on bimetallic nanocatalysts. Next is Xi Zhang, from the Key Laboratory of Organic Optoelectronics and Molecular Engineering, who has published six articles (WFC = 5.2) on supramolecules, followed by Jinghong Li, from the department of chemistry, with seven (WFC =

BEIJING ANALYSIS

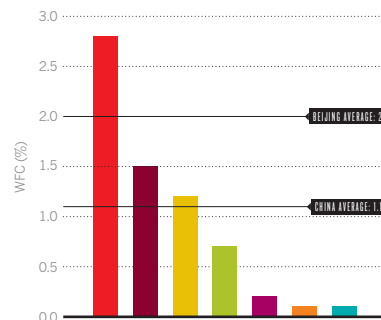
Collaboration rate

Institutes in Beijing are highly collaborative, led by UCAS.



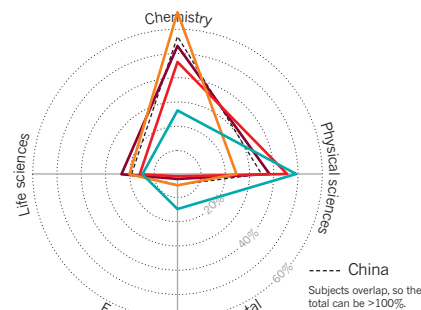
Nature and Science ratio

Tsinghua University has the highest proportion of papers in these two journals.



Institutional subject spread

Beijing Normal University has the most balanced spread of subject areas.



Shanghai

Shanghai has long been the commercial and financial centre of China. Because of its leading life–science research institutions, the city has become the hub for multinational pharmaceutical companies establishing a presence in China.

ARTICLE COUNT (AC): **1,646**

FRACTIONAL COUNT (FC): **734**

WEIGHTED FRACTIONAL COUNT (WFC): **712**

Shanghai, the most populous city in China, has undergone rapid expansion and economic transformation over the past few decades. Electronics, car manufacturing and steelmaking have long been integral industries, and now fine chemicals and biopharmaceuticals are becoming significant. In 2013, the metropolis spent US\$12 billion (3.4% of its gross domestic product, GDP) on research, and signed 9,274 technology transfer agreements — 86% of which were for electronic data services, biopharmaceuticals and advanced materials. Innovation-based industries are now responsible for 40% of the city's GDP. Only one Chinese city — Beijing — has a higher output in the Nature Index.

Shanghai is home to 68 universities, 58 research institutes, 328 hospitals, and 400 joint venture or foreign-owned research centres. Of these, 63 institutions (including 13 institutes of the Chinese Academy of Sciences, CAS) are represented in the Nature Index. The Zhangjiang Hi-Tech Park, located in the central district of Pudong, is home to dozens of multinational pharmaceutical companies, including GSK, Roche, Novartis and Pfizer. This concentration of expertise cements Shanghai's position as the world's fastest-growing city in terms of economic contribution to the life-sciences industry, according to a 2012 survey conducted by the

Swiss consultancy BAK Basel Economics.

Nearly one-third of Shanghai's research output is in the life sciences — a greater proportion than the national average (see 'City subject spread'). Overall, the major contributing institutions are Fudan University, Shanghai Jiao Tong University (SJTU), East China Normal University (ECNU), East China University of Science and Technology (ECUST), Tongji University, and the Shanghai Institutes for Biological Sciences (part of CAS, see page S56), each of which contributes between 6% and 18% of the city's weighted fractional count (WFC) — a measure of the relative contribution of an institution to the papers it has published (see 'City WFC breakdown').

NEARLY ONE THIRD OF SHANGHAI'S RESEARCH OUTPUT IS IN THE LIFE SCIENCES

Fudan is Shanghai's premier institution for higher education and has the highest WFC of any institution in the city. In 2013, the 109-year-old establishment published 255 articles (WFC = 129.2), including three (WFC = 0.8) in *Nature* and *Science* (see 'Nature and Science ratio').

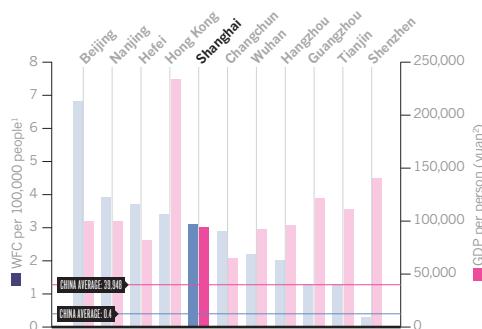
Fudan is strongest in chemistry (see 'Institutional subject spread'), particularly in materials chemistry. Huisheng Peng from the department of chemistry is the largest contributor, with nine articles (WFC = 8.6) representing more than 10% of Fudan's output in this field. Peng has developed composite nanofibres that can be woven into paper-thin capacitors or used in flexible lithium batteries. "These materials perform like conventional planar batteries but are flexible and wearable," says Peng. Such batteries might one day be used to power electronics in jackets and clothes, he adds.

Other major contributors from the same department include Dongyuan Zhao, Zhongsheng Wang and Yuping Wu. Each of these researchers published between three and five articles in a range of chemistry journals. Notably, one of Wu's papers in the journal *Nano Letters*, titled 'LiMn₂O₄ nanotube as cathode material of second-level charge capability for aqueous rechargeable batteries', was listed as one of China's 100 most influential academic papers in 2013 by the Institute of Scientific and Technical Information of China. The highest individual WFCs in life sciences at Fudan are from Yanhui Xu and Qunying Lei, both from the department of biochemistry and molecular biology. Xu contributed to three articles (WFC = 2.5) — including one in *Cell* — on crystal structures of important proteins. Lei's

SHANGHAI ANALYSIS

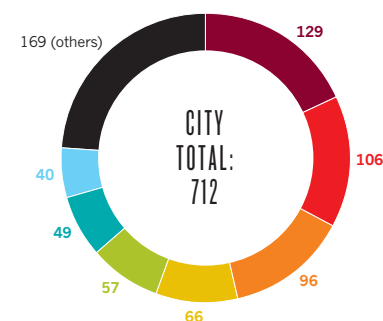
Shanghai data

Because of Shanghai's large population, both WFC and GDP rates are diluted



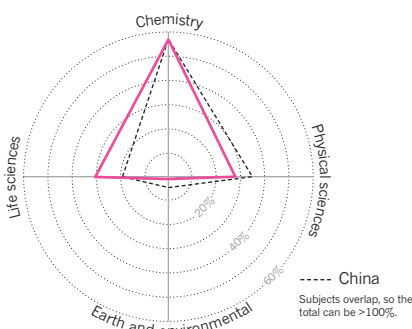
City WFC breakdown

Fudan University is top of Shanghai's 63 research institutions in the index, including many CAS institutes.



City subject spread

Shanghai is one of China's stronger cities in the index for the life sciences.



■ Fudan University ■ Shanghai Institute of Organic Chemistry* ■ Shanghai Jiao Tong University ■ East China Normal University ■ East China University of Science and Technology



The dizzying heights of the Guanhua Twin Towers at high-flying Fudan University

three articles (WFC = 2.3) were on the molecular mechanisms behind several biological processes, including lipid biosynthesis, tumour growth and cancer development.

SJTU is a comprehensive university with a 117-year history. The index shows that in 2013, the institution had a wide range of research across chemistry, physical sciences and life sciences. There are four articles in *Science*, but because of their collaborative nature, these only earned SJTU a WFC of 0.2. Indeed, the institution also stands out as the Shanghai university most open to collaboration: its AC/FC ratio is the highest among the city's ten top contributors (see 'Collaboration rate').

Wanbin Zhang from the school of chemistry and chemical engineering is SJTU's largest contributor in chemistry. He wrote six articles (WFC = 5.7) on the development of catalysts for use in asymmetric synthesis. "We discovered one of the best catalysts for promoting the hydrogenation of pentacyclic compounds," says Zhang. Pentacyclic compounds are important precursors for the synthesis of natural products and pharmaceuticals, he explains.

Other major contributors from the same school include Yong Cui, who wrote four articles with SJTU colleagues (WFC = 4) on the development of porous materials for separating chiral molecules. Shun'ai Che also co-authored

four articles (WFC = 3.7) on inorganic materials exhibiting optical activity, including one in *Nature Communications*. Che explains that, prior to her research, the only materials known to perform optical rotation for linearly polarized light were organic polymers. "We discovered that titanium dioxide is an inorganic material exhibiting this type of optical activity," she says. The advantage of inorganic material is that it can more easily be incorporated into devices made of metals or into semiconductors, she adds.

In the physical sciences, Zhengming Sheng from the Ministry of Education Key Laboratory for Laser Plasma has four articles in the index (WFC = 1.9) on laser wakefields, a technique for accelerating charged particles to high energies. But Chong Lei from the department of physics and astronomy, had a higher WFC (2), with two articles co-authored with colleagues from the same university on tiny sensors for detecting microbeads and antigens.

Life sciences comprise just over a quarter of SJTU's output — one of the highest proportions among Shanghai universities. In this realm, the most prolific researcher is Saijuan Chen from the Shanghai Center for Systems Biomedicine. Chen, who researches leukaemia, published four articles (WFC = 2.6) in 2013 in *PNAS*. Other major contributors in the life sciences include Guang Ning from the laboratory of

endocrine and metabolic diseases, who published two articles (WFC = 1.4): one in *Nature Cell Biology* on white-to-brown fat transition; and one in *Nature Communications* on a special class of pancreatic tumours. Dabing Zhang, from the school of life sciences and biotechnology, also co-authored two papers (WFC = 1.3) — including one in *Nature Communications* — on hybrid rice. "We discovered a novel mechanism for controlling flowering development," says Zhang. This knowledge could be useful for improving rice yield.

ECNU and ECUST are the two Shanghai universities that concentrate most on chemistry. In 2013, ECNU published 123 articles (WFC = 65.6) in the index, and derives 63% of its WFC from chemistry journals. ECUST has fewer articles overall, at 95 (WFC = 56.8), but as it is almost entirely focused on chemistry it has a higher WFC than ECNU in this field.

Wenhao Hu, from ECNU's Institute for Advanced Interdisciplinary Research in Science and Technology, is the university's most prolific contributor; he co-authored eight articles (WFC = 7.4) on organic synthesis. "We discovered several three-component reactions for synthesizing small-molecule drugs," says Hu.

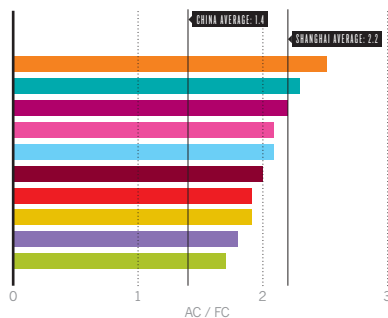
At ECUST, the most prolific contributors were Yitao Long and Huagui Yang from the school of chemistry and biomolecular engineering. Long contributed five articles (WFC = 4.1) on nanoparticles, and Yang wrote five articles (WFC = 3.6) on solar cells.

Although Tongji's output is only the fifth largest by WFC of Shanghai universities (excluding CAS institutes), it is first in terms of the proportion of publications in *Nature* and *Science*. In 2013, the 110-year-old establishment published five articles (WFC = 1.1) in these two journals, representing 2.8% of its WFC. One of these was a paper entitled 'Thin crust as evidence for depleted mantle supporting the Marion Rise' by Huaiyang Zhou — and is notable for being the first *Nature* paper in the field of marine geology with a Chinese lead author. ■

SHANGHAI ANALYSIS

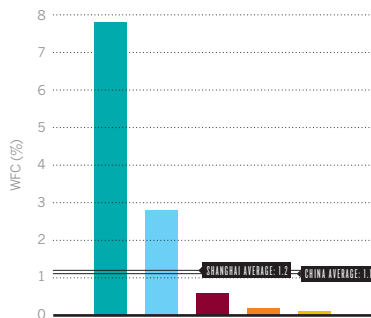
Collaboration rate

SJTU has the highest collaboration ratio of the top ten Shanghai research establishments.



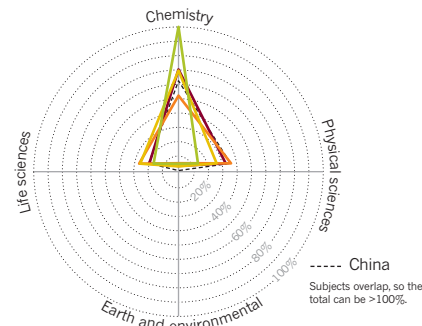
Nature and Science ratio

SIBS has one of the highest ratios for publications in *Nature* and *Science* of any Chinese institution.



Institutional subject spread

Excluding specialist CAS institutes, ECUST is the Shanghai institution most dedicated to chemistry.



*CAS institute. 1. 360doc.com; 2. cnpop.org; yuan = US\$0.16

■ Shanghai Institutes for Biological Sciences* ■ Tongji University ■ Shanghai Institute of Materials* ■ Shanghai Institute of Ceramics* ■ Shanghai Institute of Microsystem and Information Technology*

Hong Kong

Hong Kong has enjoyed 17 years of prosperity and academic freedom since the transfer of sovereignty from the UK to China. But with political unrest and increasing competition from mainland cities, it needs to rethink its long-term strategy.

ARTICLE COUNT (AC): **517**
FRACTIONAL COUNT (FC): **250**
WEIGHTED FRACTIONAL COUNT (WFC): **241**

Hong Kong, the former British colony turned special administrative region, has always been China's favoured city for science and technology (the city is ranked 10 among 143 economies worldwide in the Global Innovation Index 2014) because of its robust intellectual property protection and legal system. The city has transformed itself into an educational hub for serving the Asia-Pacific region. However, in recent years Hong Kong has faced a challenge — it must keep pace with China's scientific development or risk losing its competitive edge to nearby cities on the mainland, including Guangzhou and Shenzhen.

To give the city an edge, on 29 October 2014 the Hong Kong Legislative Council passed a resolution to establish the Innovation & Technology Bureau. The aim of the bureau is to support start-up companies and provide financial assistance for universities and research institutions to commercialize research.

While most cities in China have only one or two major contributing universities, Hong Kong has six — each comprising between 5% and 29% of the city's weighted fractional count (WFC; see 'City WFC breakdown'). In the index, WFC is a measure of the contribution of an institution to the papers its scientists have co-authored.

The University of Hong Kong (HKU) is the city's premier institution for higher education

and its largest contributor to high-quality journals — as shown by its WFC. In 2013 the 103-year-old establishment published five articles in *Nature* and *Science*, representing 1.2% of its WFC — a larger proportion than any other Hong Kong institution (see 'Nature and Science ratio'). HKU is responsible for half of Hong Kong's 10 articles in these two journals, but represents 71% of the city's WFC in this count.

“HKBU IS A SMALL UNIVERSITY BUT THE DEPARTMENTS WORK TOGETHER VERY CLOSELY”

Although it derives only 17% of its WFC from the life sciences (see 'Institutional subject spread'), HKU is strong in the field of microbiology/virology. A closer look at the index data reveals that Yi Guan and Joseph Sriyal Malik Peiris from HKU's school of public health are the most significant contributors in this area, with three articles in the index on the infectivity and transmission of avian and swine influenza viruses (including one in *Nature* and one in *Science*). “We have identified the source and

provided a detailed assessment on the infectivity, transmissibility, and pathogenicity of H7N9 influenza viruses,” says Guan. An avian flu virus, H7N9 caused 130 human infections and 40 deaths in China in March 2013. “Our work is thus far the most comprehensive piece of research on H7N9 influenza viruses.”

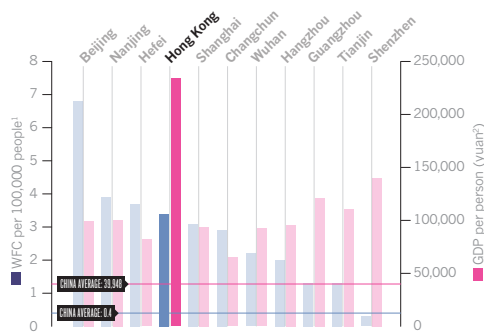
HKU also has the highest WFC of any Hong Kong institution in the physical sciences, having published 70 articles in this area (WFC = 26) mostly concerning advanced materials. Prolific researchers in the index include Wallace Chik Ho Choy at the department of electrical and electronic engineering, with four articles on organic solar cells, and Shunqing Shen and Haizhou Lu from the department of physics, who published three articles on the quantum properties of topological insulators — novel materials whose interior behaves like an insulator but whose exterior behaves like a conductor.

The Hong Kong University of Science and Technology (HKUST) has Hong Kong's second highest WFC in the physical sciences, with 49 articles in this field (WFC = 22). Two researchers from the department of physics are responsible for many of these publications. Ping Sheng is the largest contributor, with four articles in the Nature Index on graphene and metamaterials (engineered materials with optical properties not found in nature). The second is Penger Tong,

HONG KONG ANALYSIS

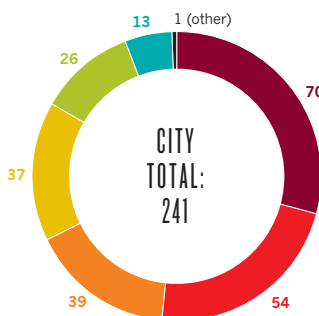
Hong Kong data

Hong Kong has the third highest GDP of any Chinese city, but the highest GDP per person.



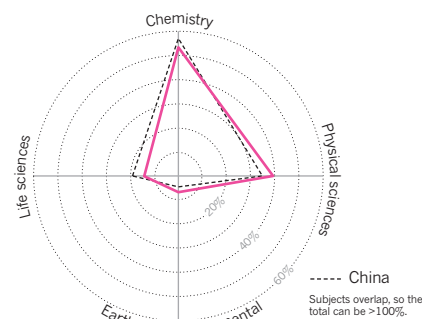
City WFC breakdown

Hong Kong has six major contributing institutions.



City subject spread

The distribution of Hong Kong's research is similar to that of China's overall.





Not yet 25 years old, yet HKUST is challenging much older institutions

who published three articles on colloidal monolayers, a model system for studying the structure and dynamics of complex fluids.

“This year we have developed acoustic metamaterials that can absorb low-frequency sound,” says Sheng. He adds that these types of materials will be useful in soundproofing homes and music studios from environmental noise.

The Chinese University of Hong Kong (CUHK) is also heavily focused on a range of physical sciences, and has many papers in the index published wholly by in-house researchers. The work of four researchers stands out from the data in 2013: Daniel Hock Chun Ong from the department of physics published two articles on the direct imaging of surface plasmon polaritons, which have important implications for Raman spectroscopy and hence molecular identification. Jianbin Xu from the department of electronic engineering published one article in *Nature Photonics* on graphene-based photodetectors with high responsivity, which increases the wavelengths of light that can be detected and hence widens the range of applications for such sensors. Qian Miao from the department of chemistry published four wholly authored articles on the synthesis of organic materials, while Zuowei Xie from the State Key Laboratory of Synthetic Chemistry published four articles on the preparation of derivatives of

carborane (a cluster composed of boron, carbon and hydrogen atoms).

Hong Kong Polytechnic University (Poly U) is the city’s institution most focused on geosciences, which account for 9% of its WFC. In 2013, it published five articles in this field, including astrogeologist Bo Wu’s landmark paper in *Earth Planetary Science Letters* on lunar topographic models. In physical sciences, which make up more than a third of Poly U’s output, the largest contributor to the Nature Index journals is Jianhua Hao from the department of applied physics. Working alone, Hao wrote three articles on functional thin films and heterostructures (all in *Applied Physics Letters*).

Poly U stands out in another measure in the Nature Index as the Hong Kong institution that has collaborated most actively, with an AC/FC ratio higher than the city’s other major universities.

City University of Hong Kong (City U) shows its strengths in physical sciences, with 44 articles in the index in this field — mainly from its department of physics and materials science. And it is in the materials science subset where it is particularly strong. Three researchers are responsible for the majority of these publications, led by Wenjun Zhang, who published four articles on nanowires and graphene. “Our materials help enhance the signals from surface-enhanced

Raman spectroscopy, a technique often used in bioimaging and medicine,” says Zhang.

The other two major contributors are Jensen Tsan Hang Li and Johnny Ho. Li has three articles on metamaterials and transformation optics, and Ho published three articles on the electronic properties of nanowires. “Our expertise has been traditionally in materials science, but in the future we would like to broaden our scope,” says Jian Lu, who is also vice-president of research and Technology at City U.

Of the Hong Kong institutions in the index, Hong Kong Baptist University (HKBU) has the highest proportion (66%) of publications in chemistry — well above the national average. Ricky Man Shing Wong from HKBU’s Institute of Advanced Materials is the largest contributor by WFC, with two articles on the development of fluorescent probes (WFC = 1.8). “We created an efficient multi-photon system for turning red light blue,” says Wong. “Such systems can serve as high-energy coherent sources for use in lasers and imaging applications.”

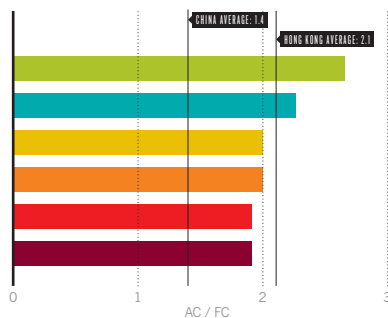
Edmond Dik Lung Ma from the department of chemistry published three articles (WFC = 1.7) in the related field of luminescent probes; work that came from an internal collaboration. “HKBU is a small university but the departments work together very closely,” he says. In 2013, Ma teamed up with researchers from the School of Chinese Medicine to detect proteins, measure enzyme activities and screen novel inhibitors. This collaboration ultimately led to the discovery of novel metal complexes for treating skin cancer.

In 2011, HKBU established the Institute of Creativity in order to enhance interdisciplinary research and academic exchange. Chemist Raymond Wai Yeung Wong, associate director of this new institute, says it has helped him receive valuable advice from colleagues outside his field to help his research. Wong has six articles (WFC = 1.6) in the index covering heterometallic complexes, which are used in highly efficient organic solar cells and light-emitting diodes. ■

HONG KONG ANALYSIS

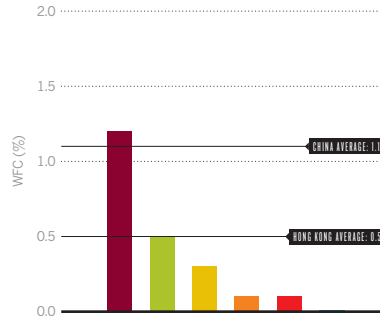
Collaboration rate

Hong Kong Polytechnic University has the highest collaboration rate.



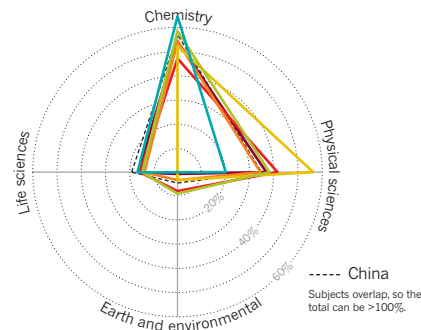
Nature and Science ratio

Only HKU exceeds the national average for papers in these two journals.



Institutional subject spread

Chemistry and physics are preferred across all institutions.



Nanjing

Nanjing has long been a hotbed for scientific discovery and technological innovations. Now, through promoting materials science and astrophysics, the former capital city hopes to step out from the shadows of its neighbour and rival Shanghai.

ARTICLE COUNT (AC): **795**
FRACTIONAL COUNT (FC): **362**
WEIGHTED FRACTIONAL COUNT (WFC): **310**

Nanjing, surrounded by green mountains and rivers, is the capital city of Jiangsu province. Though it has long been a popular destination for tourists, the 'ancient capital' is often overlooked by foreign investors who flock to nearby, and much larger, Shanghai.

Nanjing's 2012 budget for scientific research and development was US\$1.5 billion, comprised of equal contributions from local government and industry. The city's eight pillar industries in the high-tech sector are supported by more than 100 universities and research institutions, including the premier institution for education — Nanjing University (NJU).

NJU is by far Nanjing's largest contributor to the Nature Index, and fourth overall in China by weighted fractional count (WFC). In 2013, the university published 391 articles (WFC = 194.6), accounting for 64% of the city's WFC. Despite this output, NJU contributed only one article (WFC = 0.05) to *Science* and none to *Nature*.

NJU derives most of its WFC from chemistry. Huangxian Ju, Jingjuan Xu and Hongyuan Chen from the school of chemistry and chemical engineering are NJU's largest contributors. Ju — also the director of the Ministry of Education Key Laboratory of Analytical Chemistry for Life Sciences — produced 13 articles (WFC = 11.8) on fluorescent sensors, which have applications in bioimaging. Xu and Chen co-authored nine

articles (WFC = 8.4) on electrochemiluminescence, a biosensing technology for detecting cell surface proteins and DNA.

NJU is also productive in astrophysics, which makes up 15% of its fractional count (FC). However, owing to the down-weighting of astrophysics journals in the index, these publications contribute a WFC of only 6.9. Jilin Zhou and Zigao Dai from the school of astronomy and space science contribute the most to this field. Zhou co-authored four articles on planetary formation, while Dai contributed to three on gamma-ray bursts — extremely energetic explosions observed in distant galaxies.

Nanjing also has six smaller research universities and one institute of the Chinese Academy of Sciences (CAS) that each contribute 2–10% of the city's WFC. Southeast University (SEU) has the second largest output and is focused on the physical sciences, which make up two-thirds of its WFC. In 2013, the 112-year-old institution published 65 articles (WFC = 30.9), with Tiejun Cui, the vice president of the school of information science and engineering, being the most prolific contributor. Cui led 11 articles (WFC = 6.1) on transformation optics, a novel class of materials with potential use in stealth devices.

Nanjing Medical University (NJMU), founded in 1934, was one of the first institutions to offer postgraduate medical education

in China. In 2013, the university published 28 articles (WFC = 7.3), with life sciences research representing 81% of its output. It also has the highest proportion of papers in *Nature* and *Science* in the city, which comprise 4.6% of its WFC. NJMU's president, Hongbing Shen, is the most active contributor to the index, having led five genome-wide association studies (WFC = 1.4), all published in *Nature Genetics*.

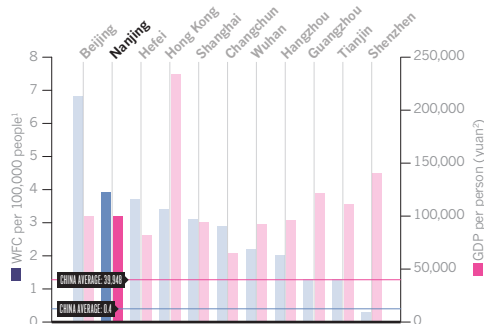
Nanjing is also the strongest city in China for astrophysics, which comprises 18% of its FC — ahead of Beijing (11%), Hefei (8%) and Hong Kong (5%). This knowledge base is largely due to the Purple Mountain Observatory (PMO), a CAS institute. Last year, PMO published 102 articles (WFC = 6.3), the majority from three contributors. Dejin Wu, the deputy director of the division of dark matter and space astronomy, published six articles on solar flares and coronal loops; Yizhong Fan published five articles on dark matter and gamma-ray bursts; and Xuefeng Wu contributed to seven articles on gamma-ray bursts.

PMO has a number of high-profile projects underway. "We are in the preparation stage of launching our own observation satellites into space, and a team of scientists will also be setting up an observatory in Antarctica," says Xuefeng Wu. "China's research capabilities in astrophysics have come a long way since the 1980s." ■

NANJING ANALYSIS

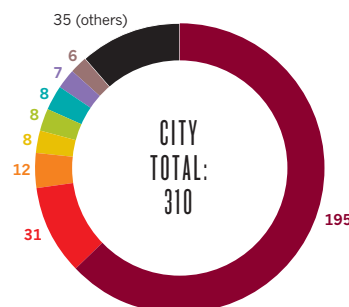
Nanjing data

Nanjing is second only to Beijing in terms of WFC per person, and in the top five for GDP per person.



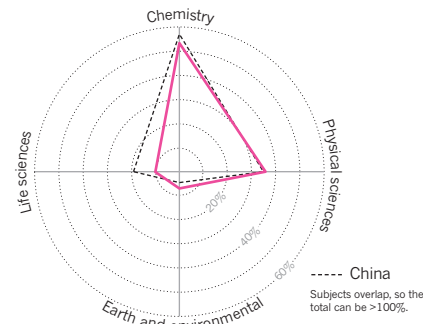
City WFC breakdown

Nanjing University dominates the city in terms of scientific output in the index.



City subject spread

The majority of Nanjing's research is in chemistry.



■ Nanjing University ■ Southeast University ■ Nanjing University of Technology ■ Nanjing University of Posts and Telecommunications
■ Nanjing University of Information Science and Technology ■ Nanjing Normal University ■ Nanjing Medical University ■ Purple Mountain Observatory*

*CAS institute
1. 360doc.com; 2. cnpop.org; yuan = US\$0.16

Wuhan

Wuhan is the booming capital of the eastern inland province of Hubei. The city is investing heavily in research and development and has become China's 'optics valley'.

ARTICLE COUNT (AC): **486**
FRACTIONAL COUNT (FC): **222**
WEIGHTED FRACTIONAL COUNT (WFC): **217**

Wuhan is an important centre for manufacturing, information technology, transportation and education. The optoelectronics industry in particular has enjoyed rapid growth in recent years; in 2013, Wuhan produced more than 127 million kilometres of fibre optic cables, the largest output of any Chinese city.

In November 2014, Wuhan initiated the eighth instalment of its highly selective '3551 optics valley talent program'. Launched in 2011, the programme aims to recruit global leaders and young researchers to work in five designated areas: information technology, biotechnology, energy and environment, specialized equipment, and a sector that China calls modern service (essentially, IT-enabled service industries). The latest instalment increases the funding available for top researchers to US\$16.3 million.

Of Wuhan's higher education institutions, Wuhan University (WHU) and Huazhong University of Science and Technology (HUST) are the two largest in the index. Together they account for two-thirds of the city's weighted fractional count (WFC) — a measure of the relative contribution of an institution to the papers it has published. WHU has 154 articles (WFC = 98.8) in the index. The 121-year-old establishment is strongest in chemistry, which accounts for 64% of its WFC. Aiwen Lei from the department of

chemistry was WHU's most prolific chemistry researcher by some distance, having led 20 articles (WFC = 16.4) on the use of free radical coupling reactions in organic synthesis. Lei explains that the reactions can be used to introduce various functional groups onto organic molecules. "The free radicals bind by forming carbon-carbon bonds and releasing hydrogen gas."

Hongbing Shu from the college of life sciences was the largest contributor in the life sciences, having published four articles (WFC = 3.4) on innate immunity. Specifically, Shu studied how enzymes suppress signalling molecules, including tumour necrosis factors and interleukins. The findings have important implications in the development of cancer treatments and for understanding inflammation. Xiangdong Fu from the same college is also a significant contributor to the index, with four articles (WFC = 1.8) — including two in *Cell* — on cell reprogramming. One of his articles, "Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated RNA circuits", was listed as one of China's 100 most influential academic papers in 2013 by the Institute of Scientific and Technical Information of China.

HUST's research strengths lie in physics and material sciences. Last year, the university had 109 articles (WFC = 43.6) in the index, of which 61 (WFC = 25.3) were in this field. Xiangshui

Miao and Jingping Xu from the school of optical and electronic information are HUST's largest contributors. Miao published four articles (WFC = 2.9) on ultrathin films, while Xu published two (WFC = 1.5) on metal oxide semiconductors.

Wuhan has four other national key universities in the index: Central China Normal University (CCNU), China University of Geosciences (CUG), Huazhong Agricultural University and Wuhan University of Technology. Each of these institutions contributed between 4% and 7% of the city's WFC.

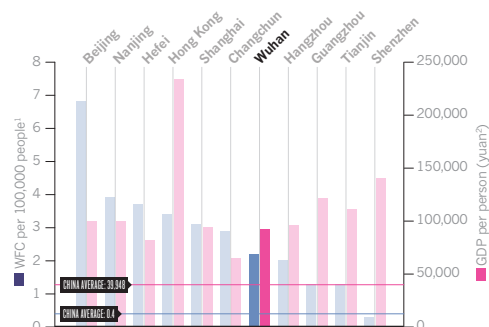
Of CCNU's 43 articles (WFC = 15.9), 84% were in chemistry. The largest contributor was Anxin Wu from the department of chemistry, with four papers (WFC = 3.9) on drug design and development. "We have made much progress in the synthesis of natural products," says Wu. He adds that 40 natural compounds were made through one-pot synthesis — multiple reactions in a single reactor.

CUG is Wuhan's largest contributor to earth and environmental sciences. This highly specialized institution has 65% of its WFC in this field. The work was broadly shared, with no CUG researcher contributing to more than one article in the index. CUG does have one wholly authored paper (WFC = 1) in *Earth and Planetary Science Letters*, led by Yongfeng Wang from the department of geology. ■

WUHAN ANALYSIS

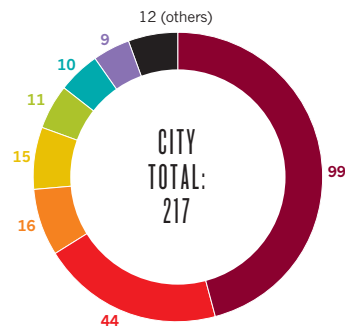
Wuhan data

Wuhan's large population dilutes its ratios of gross domestic product (GDP) and WFC per person.



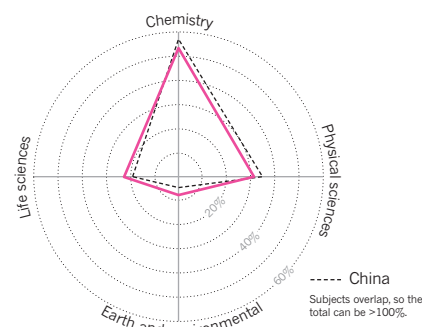
City WFC breakdown

Wuhan has 33 institutions (including CAS) in the index, the third largest after Beijing and Shanghai



City subject spread

Wuhan's subject spread is similar to China's overall.



■ Wuhan University ■ Huazhong University of Science and Technology ■ Central China Normal University ■ China University of Geosciences
■ Huazhong Agricultural University ■ Wuhan Institute of Physics and Mathematics* ■ Wuhan University of Technology

*CAS institute; 1. 360doc.com; 2. cnpop.org; yuan = US\$0.16

Hefei

Hefei has kept a low profile for many years. However, as the University of Science and Technology of China continues to break new ground in the physical sciences, the city is moving into the global spotlight.

ARTICLE COUNT (AC): **530**

FRACTIONAL COUNT (FC): **226**

WEIGHTED FRACTIONAL COUNT (WFC): **212**

Hefei, the capital city of the eastern Anhui province, is the smallest of the cities profiled in this supplement. But it is growing rapidly in population, disposable income levels and gross domestic product (GDP) — and when it comes to the pursuit of basic science, it holds its own among much larger cities.

Hefei is home to the University of Science and Technology of China (USTC), one of three universities affiliated with the Chinese Academy of Sciences (CAS). When USTC's first president Moruo Guo laid its foundation stone in 1958, he set out its mission to focus on basic research and to nurture world-class talent. Ever since, the institution has been faithful to this ideal.

USTC is by far Hefei's largest contributor to the Nature Index, and fifth overall in China by weighted fractional count (WFC) — a measure of the relative contribution of an institution to the papers it has published. In 2013, USTC accounted for 83% of Hefei's WFC (see 'City WFC breakdown'). In addition, USTC made a strong showing in *Nature* and *Science* journals, with eight articles (WFC = 1.7) representing 1% of its WFC — well above other Hefei institutions.

Most of USTC's WFC is in chemistry. Yi Xie from the division of nanomaterials and nanochemistry is the university's most prolific

researcher in this field, with 17 articles (WFC = 13.2) in the index, mostly on graphene-like materials. "Graphene comprises carbon atoms only, so its structure and chemical properties are rather simple," she says. "We are developing graphene-like inorganic materials with unusual properties that may find applications in photocatalysis and biomedicine." Another successful chemistry researcher from the same division, Shuhong Yu co-authored 12 articles (WFC = 11.4) in 2013 on aerogels. His team manufactured carbon aerogels by freeze-drying bacterial cellulose and attaching CH groups, Yu explains. The resultant hydrophobic material "can be used to remove organic pollutants from water," he adds.

However, it is physics for which USTC is most renowned. The strongest contributor in this field is Guangcan Guo from the Key Laboratory of Quantum Information. Guo has 11 articles (WFC = 8.7) on quantum optics, quantum communication and topological superfluids in the index. His team developed a technique to record the orbital angular momentum of a single photon. An advance that Guo says "represents an important first step towards the realization of long-distance quantum communication." USTC's president, Jianguo Hou, from the division of atomic and molecular sciences, is another active researcher. In 2013, Hou contributed to six

articles (WFC = 5.73) — including one in *Nature* — about molecular and surface imaging.

USTC also hosts the CAS Key Laboratory for Research in Galaxies and Cosmology in its department of physics. Tinggui Wang is the biggest contributor in astrophysics, with seven articles on quasars and active galactic nuclei. However, owing to the down-weighting of astrophysics journals in the index, these papers only add a WFC of 1 to the institution (see 'A guide to the Nature Index', page S76).

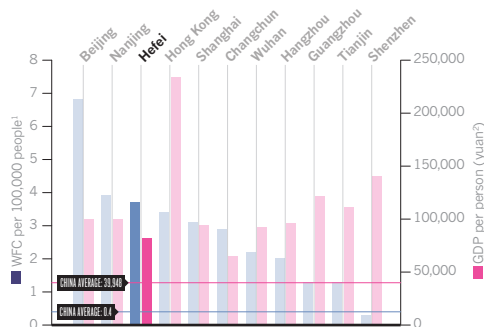
Life-science research comprises only 11% of USTC's output. Much of the university's best research in this field is carried out by Zhigang Tian from the division of structure and function of biomacromolecules. Tian's six articles (WFC = 4.2) on the regulatory functions of natural killer cells represent more than 13% of USTC's total life science WFC.

Hefei is also home to Hefei University of Technology (HFUT), an older but smaller institution with a focus on engineering. Linbao Luo from the laboratory of micro/nano functional materials and devices and Ruzhong Zuo from the school of materials science and engineering are HFUT's top two contributors, both publishing almost exclusively in *Applied Physics Letters*. Luo led two papers on nanowires — both authored wholly in-house — while Zuo contributed to two papers (WFC = 1.7) on lead-free ceramics. ■

HEFEI ANALYSIS

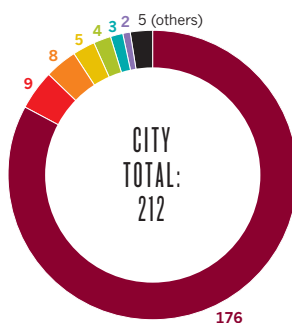
Hefei data

Hefei's WFC per person is the third highest in the index, indicating a strong concentration of research.



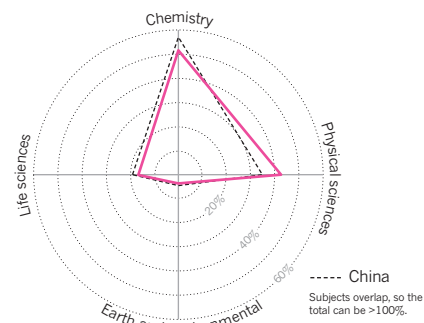
City WFC breakdown

University of Science and Technology of China is by far Hefei's largest contributing institution.



City subject spread

Hefei is focused on the broad range of physical sciences.



■ University of Science and Technology of China ■ Hefei University of Technology ■ Institute of Solid State Physics*
■ Institute of Intelligent Machines* ■ Anhui University ■ Institute of Plasma Physics* ■ Anhui Medical University

*CAS institute: 1. 360doc.com; 2. cnpop.org; yuan = US\$0.16

Hangzhou

Hangzhou, a tourist hotspot best known for its historical relics and natural scenery, has long been a source of inspiration for Chinese artists. Today the city is home to a new generation working at the intersection of science and e-commerce.

ARTICLE COUNT (AC): **377**
FRACTIONAL COUNT (FC): **178**
WEIGHTED FRACTIONAL COUNT (WFC): **178**

Hangzhou, an important manufacturing base and East China's regional logistic hub, is fast becoming the world's largest e-commerce centre. The city is home to the Alibaba Group, China's leading e-commerce service provider with more than 300 million customers and an estimated market value of US\$231 billion. In September 2014, Alibaba raised US\$25 billion from its initial public offering (IPO), making it the largest IPO in US history.

The presence of this commercial giant is shaping local infrastructure. In 2008, Alibaba and Hangzhou Normal University (HZNU) co-founded Alibaba Business College, a centre for education and training on e-commerce, data mining and modern logistics. And in 2013, they established the Alibaba Research Center for Complexity Sciences for research into econometrics and the physics of complex systems. The college has already published several papers in scientific journals.

Hangzhou's weighted fractional count (WFC) and fractional count (FC) are the same, which shows the city has no astrophysical research. It does, however, have several institutions engaged in other areas of physical sciences. The most famous is Zhejiang University (ZJU), which is the city's largest contributor to the Nature Index, and

sixth overall in China by WFC — a measure of the relative contribution of an institution to the papers it has published. In 2013, ZJU published 289 articles (WFC = 150.4), accounting for 85% of the city's WFC (see 'City WFC breakdown'). In addition, ZJU has a strong representation in *Nature* and *Science* journals, with seven articles (WFC = 1.7) representing 1.1% of its WFC — well above other Hangzhou institutions.

ZJU is strong in both chemistry and the physical sciences. Feihe Huang from the department of chemistry was the largest contributor in this field, with a total of 14 papers (WFC = 13.1) in a range of journals. Last year in particular, his paper "A supramolecular cross-linked conjugated polymer network for multiple fluorescent sensing" — published in *Journal of the American Chemical Society* — was listed by the Institute of Scientific and Technical Information of China as one of 'China's top 100 most influential academic papers' in 2013. "We have developed a special polymer that fluoresces in the presence of ammonia," says Huang. The technology can be used to detect gas leaks in refrigeration systems among other applications, he explains.

Chao Gao from ZJU's department of physics was the largest contributor in the physical sciences. Last year, Gao published three

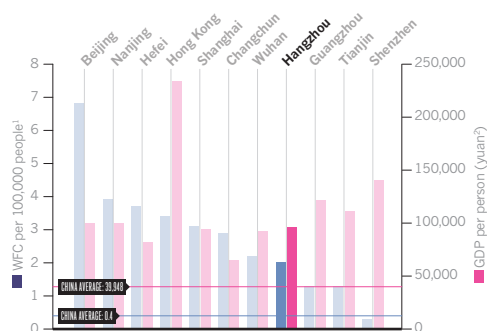
articles (WFC = 3) on graphene and carbon aerogels in the journal *Advanced Materials*. "Our carbon aerogel, with a density of 0.16 mg/cm, is currently the lightest material in the world," says Gao, adding that this ultra-light, porous, synthetic material has potential applications in thermal insulation, oil adsorption and gas sensing.

Hangzhou is also home to HZNU, a smaller and younger university specializing in education, literature and mathematics. The institution contributes approximately 5% of the city's WFC. According to the index, HZNU is strong in both physics and chemistry. Zhi-fang Li from the laboratory of organosilicon chemistry and material technology is HZNU's largest contributor in chemistry. Li led two articles (WFC = 2) on silylenes — highly reactive intermediates to which a broad range of functional groups can be added. Zujin Zhao from State Key Laboratory of Luminescent Materials and Devices is another top contributor, with two articles in the index (WFC = 1) on novel luminescent materials. "We showed that by decorating a tetraphenylethene core with four aromatic groups, the material displays enhanced emission and fluorescence efficiencies," says Zhao. The finding has implications for the development of organic light-emitting diodes. ■

HANGZHOU ANALYSIS

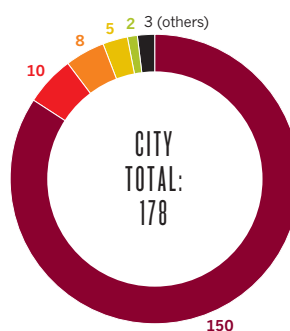
Hangzhou data

Hangzhou's relative prosperity doesn't translate into a high WFC per person.



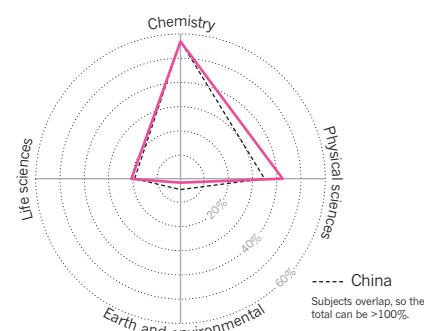
City WFC breakdown

Zhejiang University is Hangzhou's dominant research institution in the index.



City subject spread

Hangzhou's subject spread is similar to China's with a focus on physical sciences.



Changchun

Changchun has traditionally been a manufacturing centre, producing goods ranging from cars to processed food. In line with China's growth and reform, the city is diversifying its economy by leveraging its research base.

ARTICLE COUNT (AC): **394**

FRACTIONAL COUNT (FC): **226**

WEIGHTED FRACTIONAL COUNT (WFC): **224**

Changchun, the capital and largest city of the northeastern province of Jilin, is home to many production industries. Its government estimates that output reached US\$150 billion in 2013, of which well over 90% came from the manufacture of cars, food, bio pharmaceuticals and construction materials.

However, Changchun wants to become an innovation-based economy. On 21 January 2014, the provincial governor announced the city would invest heavily in four new areas — photonics, chemical engineering, biochemistry and fine chemicals — and build on research from its four major research institutions: Jilin University (JLU); Northeast Normal University (NENU); and the two institutes of the Chinese Academy of Sciences (CAS) — the Changchun Institute of Applied Chemistry (CIAC) and the Changchun Institute of Optics, Fine Mechanics and Physics (CIOMP). Collectively, these institutions account for more than 98% of the city's weighted fractional count (WFC).

JLU is responsible for most of these papers. Jihong Yu and Guangshan Zhu from the State Key Laboratory of Inorganic Synthesis and Preparative Chemistry are the two largest contributors. In 2013, Yu published eight articles (WFC = 6.6), and Zhu seven (WFC = 6.2). Both researchers study the design and synthesis of porous materials, which have diverse applications

including carbon sequestration, water purification, catalysis and chromatography.

Myongssoo Lee, from the State Key Laboratory of Supramolecular Structure and Materials, is JLU's third highest contributor by WFC to chemistry. Lee joined JLU in 2013 but has already published three papers (WFC=2.1), on the self-assembly of nanomaterials.

CIAC is Changchun's powerhouse for high-quality chemistry research. Husband-and-wife team Xiaogang Qu and Jinsong Ren from the State Key Laboratory of Rare Earth Resource Utilization — who have been inseparable since graduating from the California Institute of Technology — have made the largest contribution by WFC. They published 20 articles (WFC = 16.3) on topics including artificial enzymes, catalysis, gene delivery and cell imaging. "We have developed novel biomimetics that could simulate cellular processes," says Ren. "We have made artificial enzymes that could serve a range of industrial applications that are cheaper and more rugged than their natural counterparts."

Another CIAC couple, Erkang Wang and Shaojun Dong from the State Key Laboratory of Electroanalytical Chemistry, are the third and fourth largest contributors with 24 articles (WFC = 16.0) on nanomaterials and G-quadruplexes (specific formations of nucleic acids). "We are working on enzymatic fuel cells that can produce

energy from bioavailable substrates," says Dong. In implants, such devices could power memory and electrical circuits in pacemakers.

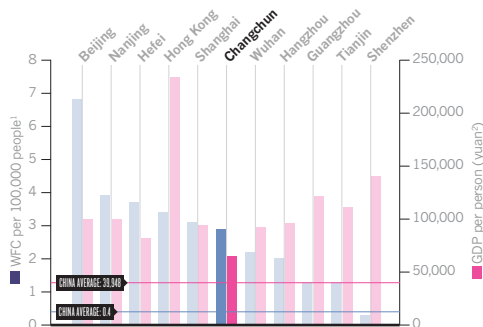
NENU is Changchun's top institution in the life sciences, which represent more than a fifth of its WFC. Notable contributors include palaeontologist Dongyu Hu, from the Ministry of Education Key Laboratory of Vegetation Ecology, who co-authored two articles — including one in *Nature* — on feathered dinosaurs. Hu sheds light on the evolution of birds and the origins of flight. It is chemistry where NENU excels. Zhongmin Su and Qian Zhang from the faculty of chemistry are its most prolific contributors. Su co-authored six articles (WFC = 5.1) on the synthesis of polyoxometalates (large metal clusters) and metal-organic frameworks, which have applications from catalysis to data storage. Zhang produced five articles (WFC = 4.9) on metal catalysts, notably on methods for introducing functional groups under mild conditions.

CAS institute CIOMP is the most dedicated to the physical sciences. Dezhen Shen and Jialong Zhao, from the State Key Laboratory of Luminescence and Applications, are the top two contributors with two articles each. Shen's are on the photocatalytic properties of titanium dioxide and manganese-doped zinc oxide, while Zhao's articles focus on the synthesis and application of zinc sulphide quantum dots. ■

CHANGCHUN ANALYSIS

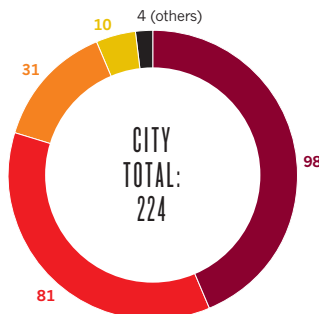
Changchun data

Per person, Changchun has the lowest gross domestic product (GDP) of the cities profiled, but an average WFC.



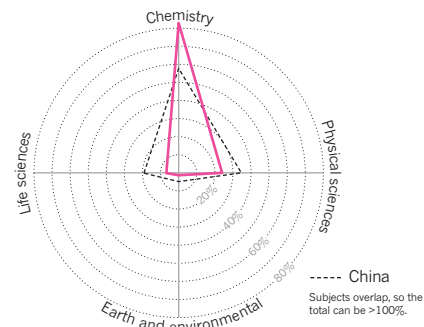
City WFC breakdown

Changchun has four major institutions, led by Jilin University.



City subject spread

Changchun is exceptionally focused on chemistry.



■ Jilin University ■ Changchun Institute of Applied Chemistry* ■ Northeast Normal University
■ Changchun Institute of Optics, Fine Mechanics and Physics*

*CAS institute: 1. 360doc.com; 2. cnpop.org; yuan = US\$0.16

Guangzhou

Prone to infectious disease outbreaks, Guangzhou this year experienced one of its worst for dengue fever. By exploiting the data from thousands of clinical cases, researchers hope to reveal new approaches for prevention and control.

ARTICLE COUNT (AC): **374**
FRACTIONAL COUNT (FC): **175**
WEIGHTED FRACTIONAL COUNT (WFC): **170**

Guangzhou, capital city of the Guangdong province and China's fourth most populous city, is a dynamic metropolis on the southern coast. It attracts more than 150 million tourists and business travellers every year. And, because of heavy human traffic and subtropical climate — compounded by the popularity of live animal markets and a local penchant for wild meat, Guangzhou regularly experiences outbreaks of infectious disease, most recently severe acute respiratory syndrome (SARS; 2002–3), H1N1 influenza (swine flu; 2009) and dengue fever (2014).

This propensity for disease provides opportunities for new research. In November 2014, the Zhongshan School of Medicine, part of Sun Yat-sen University (SYSU), held its first symposium on dengue fever control using *Wolbachia* — a bacterium that infects mosquitoes and stops the dengue virus from replicating. Researchers hope that a collaborative effort between China, Australia and the United States will lead to safe, low-cost and environmentally sound methods for eradicating the disease. Guangzhou will conduct its first field trial of the technique next year.

SYSU is Guangzhou's leader. Last year it contributed to 158 articles, accounting for 47% of the city's WFC. Qinfen Zhang co-authored an article on the protein structure of a dengue virion in *Nature Structural & Molecular Biology*.

Two-thirds of SYSU's output is in chemistry, and there are three major contributing researchers from SYSU's school of chemistry and chemical engineering. Jiepeng Zhang co-authored three articles on metal-organic frameworks, a class of porous composite materials that have wide-ranging applications from catalysis to water decontamination. Chengyong Su and Hsiuyi Chao each published two articles (WFC = 4): Su wrote about metal-organic frameworks for use in gas adsorption, while Chao's were on metal complexes for use in luminescent sensors and cell imaging. "The school has provided us with great experimental facilities, but the financial support has been limited," says Chao. "SYSU still has much to learn from the world's top universities."

From the school of physics and engineering, Biao Wang and Baojun Li each published three wholly authored articles. "We have developed fibre optic probes for use in the non-invasive control of microbes," says Li. "Our light-based technology has implications for unblocking clots and manipulating single cells in blood vessels."

South China University of Technology (SCUT) also has a strong focus on chemistry. Huanfeng Jiang from the school of chemistry and chemical engineering is responsible for half of SCUT's chemistry WFC, with 16 articles on metal-catalysed organic synthesis. From the same school, Fei Huang published two articles

on polymer solar cells. "Our polymer solar cells have high energy conversion efficiency, even at high thickness," says Huang. "Thick solar cells are a lot easier to make, so this will lower the requirements for large-scale production."

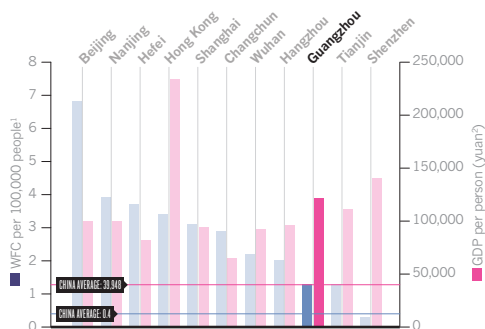
The focus of Guangzhou Institutes of Biomedicine and Health (GIBH) is medical research, so its index output is split between chemistry and life science. It also has the lowest ratio of AC to FC, indicating that many of its papers are authored by its own scientists. Duanqing Pei, the dean of GIBH, is the most prolific contributor, with six articles on novel techniques for reprogramming somatic cells (WFC = 6). Also notable are Qiang Zhu from GIBH's State Key Laboratory of Respiratory Diseases, who published four articles on organic synthesis, and Lingwen Zeng from the Key Laboratory of Regenerative Biology, also with four articles, on biosensors.

The South China Sea Institute of Oceanology (SCSIO) is a Chinese Academy of Sciences institute devoted to marine research. Ten of its 18 articles are in earth and environmental sciences, representing 43% of the city's output in this field. Jianhua Ju from the Key Laboratory of Tropical Marine Bio-resources and Ecology was the lead author on three, about the biosynthesis of marine alkaloids by bacteria. "The deep sea is full of undiscovered metabolites that can be used against antibiotic-resistant bacteria," he says. ■

GUANGZHOU ANALYSIS

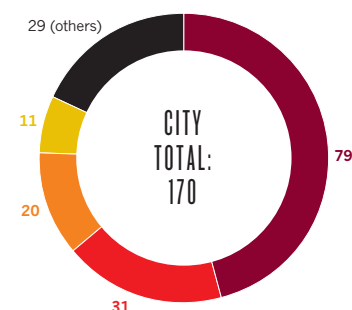
Guangzhou data

Per person, Guangzhou has one of China's highest rates of gross domestic product (GDP), but one of the lowest rates for WFC.



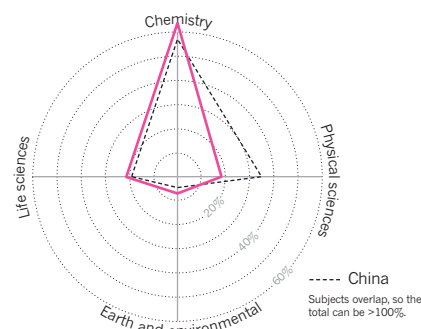
City WFC breakdown

Guangzhou has 27 institutions in the index. The largest contributor is Sun Yat-sen University.



City subject spread

Guangzhou is stronger than average in chemistry and earth and environmental sciences.



■ Sun Yat-sen University ■ South China University of Technology ■ Guangzhou Institutes of Biomedicine and Health*

■ South China Sea Institute of Oceanology*

*CAS institute
1. 360doc.com; 2. cnpop.org; yuan = US\$0.16

Tianjin

Tianjin, a major transport hub 120 kilometres southeast of Beijing, is one of four municipalities under the direct administration of central government. By leveraging the innovation of its top universities, the city hopes to lead the nation in two emerging research areas.

ARTICLE COUNT (AC): **341**

FRACTIONAL COUNT (FC): **169**

WEIGHTED FRACTIONAL COUNT (WFC): **168**

Over the last four years, Tianjin has positioned itself as an incubator for innovation. It spent US\$7 billion — or 2.98% of its GDP — on research and development in 2013 — a percentage surpassed only by Beijing and Shanghai. Through the provision of start-up subsidies, the city has increased its number of innovation-based companies to 50,000, which now account for 20% of its small- and medium-sized enterprises.

In January 2014, the Tianjin Co-Innovation Center for Chemical Science and Engineering — established by the city's two leading universities of Nankai (NKU) and Tianjin (TJU) — held a meeting of 70 distinguished scientists to explore better ways to leverage its discoveries. As a result, the centre will focus on two research areas, advanced functional materials and renewable energy. The goal will be to streamline the product development cycle, from research to patenting to technology transfer. Both universities have published work in these fields in 2013, including two articles (WFC = 2) on solar cells for NKU; and four articles on technologies including hydrogen production, solar cells and microbial fuel cells (total WFC = 2.24) for TJU.

Tianjin is strong in chemistry and though both leading universities focus on this field, NKU is historically strong in basic research, while TJU is better known for applied research.

Founded in 1919, NKU is Tianjin's premier higher education institution and the city's largest contributor to the index. In 2013, the university published 190 articles (WFC = 113.8), accounting for 68% of the city's WFC. Three researchers from its college of chemistry are responsible for most of these publications. Xiuping Yan is the most prolific, having published 12 articles (WFC = 10.5) on metal-organic frameworks, a class of porous composite materials with applications in catalysis, sensing and separation. Yan explains that these materials are significant because of their long-lasting fluorescence.

Qilin Zhou and Jun Chen are the second and third largest contributors by WFC. Zhou led eight articles — all co-authored by NKU researchers (WFC = 8) and all pertaining to asymmetric synthesis. The highlight, he says, was a chiral catalyst with a 'turnover number' (the number of molecules a catalyst converts before it is exhausted) exceeding 4.5 million — way above the typical value of the order of one thousand. Chen, meanwhile, co-authored eight articles (WFC = 6.9) on nanomaterials.

Founded in 1895, TJU is Tianjin's second largest research institution in the index with 66 articles (WFC = 33.7), representing 20% of the city's WFC. Jinlong Gong from the school of chemical engineering and technology is the largest contributor in chemistry, with seven

articles (WFC = 4.8) — including one in *Nature Communications* — on nanomaterials.

Also notable is Jun'an Ma from the department of chemistry, who published three articles (WFC = 2.8) on the synthesis of organo-fluorine compounds. "We found a way of constructing trifluoromethyl pyrazoles," says Ma. This can be used to treat HIV or arthritis, he says.

There are eight other Tianjin universities in the Nature Index. These include Tianjin Medical University (TMU), Tianjin University of Technology (TUT) and Tianjin University of Science and Technology (TUST). Among these, TMU stands out as the most collaborative. Its AC/FC ratio is the highest among the city's six major contributing universities.

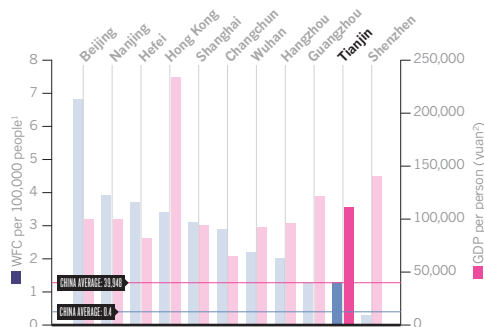
Of TUT researchers, Xianshun Zeng, in the school of materials science and engineering, is the most prolific. He co-authored three articles (WFC = 1.9) on fluorescent chemosensors which have big implications for detecting and monitoring environmental contaminants, such as palladium and bisulphate species.

TUST is the only Tianjin institution with earth and environmental science research, which represents more than 41% of its WFC. The top contributor is Hao Wei, dean of the college of marine science and engineering. Wei published two articles (WFC = 1.1) on mechanisms driving interannual ocean variability. ■

TIANJIN ANALYSIS

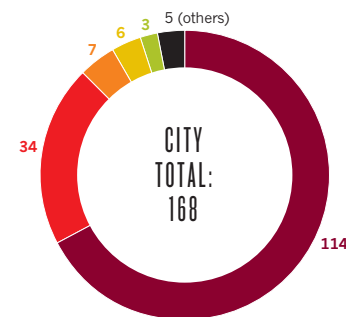
Tianjin data

Tianjin has a high rate of GDP per person, but one of the lowest rates of WFC of the cities profiled.



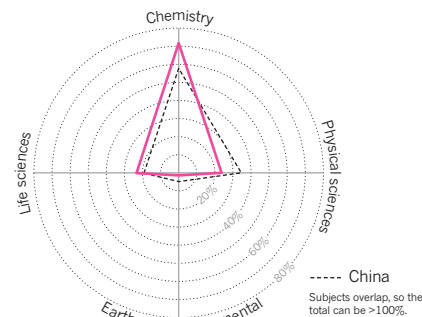
City WFC breakdown

Of Tianjin's 19 institutions in the index, Nankai is dominant by WFC.



City subject spread

Nearly three-quarters of Tianjin's WFC is derived from chemistry, way above the national average.



Shenzhen

Shenzhen, a former fishing village, is something of a miracle in China's scientific development. It has become a dominant force in genomics, proteomics and bioinformatics, and is now heading for new frontiers.

ARTICLE COUNT (AC): **107**
FRACTIONAL COUNT (FC): **35**
WEIGHTED FRACTIONAL COUNT (WFC): **35**

Until recently, Shenzhen was best known for factories in which cheap labour churned out counterfeit goods. The city has repositioned itself as one of the world's leading centres for genetics research. Shenzhen is home to BGI Shenzhen (formerly the Beijing Genomics Institute, now known as BGI), a prolific gene-sequencing organization that accounts for 50% of global sequencing capacity. In 2013, BGI acquired Complete Genomics, a US-based bioinformatics company and its closest rival — a move that will further secure BGI's dominance in the 'omics' realm.

However, BGI's success is only one aspect of Shenzhen's transformative journey. In 2011, the Shenzhen Municipal People's Government set out its twelfth 'five-year plan' to support research and innovation within six strategic emerging industries: biotechnology, internet, renewable energy, advanced materials, cultural creativity and information technology. It is hoped that by 2015, the total output value of these industrial sectors will be US\$49 billion — which equates to 20% of the city's current GDP. For comparison, the current output value of BGI is approximately US\$16 billion.

For now, Shenzhen's research strength is still predominantly in the life sciences. The city has grown its output in chemistry, although

the Nature Index shows that its WFC for this subject is still 15% below the national average.

BGI remains the largest contributing institution, with 51 articles (WFC = 15.3) in the Nature Index, accounting for 44% of the city's WFC. Included in these are seven in *Nature* and *Science* (WFC = 1.8), representing nearly 12% of its output. This means that the gene-sequencing organization has the second highest percentage WFC in *Nature* and *Science* of all Chinese research institutions, topped only by the Institute of Vertebrate Paleontology and Paleoanthropology, of the Chinese Academy of Sciences (CAS, page S56).

Jun Wang, the founder and director of BGI, led most of these publications. Last year, he contributed to 35 articles (WFC = 11.7), the most notable of which were on the genomes of bread wheat, bats and the rock pigeon. "Comparative analysis of bat genomes provides insight into the evolution of flight and immunity" published in *Science* was listed as one of China's most influential papers of 2013.

The next two major contributors from BGI are Xun Xu and Guojie Zhang, with three articles each. Xu's three (WFC = 1) were on the genomes of domestic goats, Chinese pears and upland rice; Zhang's (WFC = 0.9) were on the genomes of soft-shell turtles, green-shell turtles and one comparing the genomes

of modern domestic horses to that of a horse from the late Pleistocene.

The remainder of Shenzhen's output in the index comes mostly from the Shenzhen Institutes of Advanced Technology (SIAT) and Shenzhen University (SZU), accounting for 32% and 14% of the city's WFC, respectively.

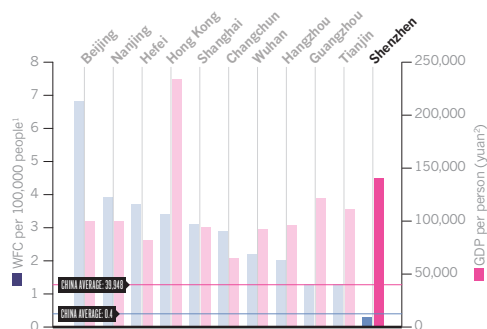
The Nature Index shows that all of SIAT's output is in chemistry. The CAS institute published 13 articles (WFC = 11.4), accounting for more than three-quarters of the city's total chemistry WFC. Most of these articles were led by analytical chemist Chunyang Zhang, who last year contributed to ten articles (WFC = 9.5) on quantum dots and amplification binding assays — tools for detecting transcription factors, enzymes and microRNAs.

Established only 31 years ago, SZU is Shenzhen's leading institution in the physical sciences. In 2013 it published seven articles in these fields (WFC = 3), most of which were in *Applied Physics Letters*. Xiacong Yuan is SZU's largest contributor, having published three articles (WFC = 0.7) on optical tweezers. "We are the first to use surface plasmon polaritons in the confinement of metal nanoparticles," says Yuan. The experiment has important implications for Raman spectroscopy, a surface imaging technique widely used in chemistry and solid-state physics. ■

SHENZHEN ANALYSIS

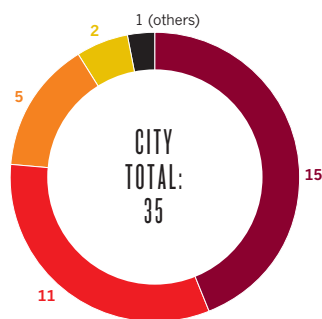
Shenzhen data

Shenzhen is the second most prosperous city after Hong Kong but lacks a comparable research base.



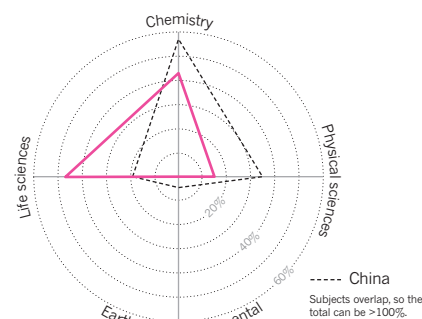
City WFC breakdown

BGI and CAS institute SIAT contribute most of the city's WFC.



City subject spread

BGI's influence means that life-science research is a substantial part of Shenzhen's WFC.



A guide to the Nature Index

A description of the terminology and methodology used in this supplement, and a guide to the functionality available online at natureindex.com.

The Nature Index is a database of author affiliations and institutional relationships, used to track contributions to articles published in a small group of highly selective journals that have been chosen by an independent group of working scientists.

Data in the Nature Index are updated monthly, with the most recent 12 months of data available under a Creative Commons licence at natureindex.com. The database is compiled by Nature Publishing Group (NPG) in collaboration with sister company Digital Science.

NATURE INDEX METRICS

There are three measures provided by the Nature Index to track affiliation data. The simplest is the article count (AC). A country or institution is given an AC of 1 for each article that has at least one author from that country or institution. This is the case whether an article has one or a hundred authors, and it means that the same article can contribute to the AC of multiple countries or institutions.

To get a better sense of a country or institution's contribution to an article, and to remove the issue of double-counting of articles, the Nature Index uses the fractional count (FC). FC takes into account the relative contribution of each author to an article. The total FC available per paper is 1, and this is shared between all authors under the assumption that each contributed equally. For instance, a paper with 10 authors means that each author receives an FC of 0.1. For authors with joint affiliations, the individual FC is then split equally between each affiliation.

The third measure is the weighted fractional count (WFC), which applies a weighting to the FC in order to adjust for the over-representation of papers from astronomy and astrophysics. The four journals in these disciplines publish about 50% of all papers in international journals in this field — approximately five-times the equivalent figures for other fields. Therefore, although the data for astronomy and astrophysics are compiled in exactly the same way as for all other disciplines, articles from these journals are assigned one-fifth the weight of other articles (i.e. the FC is multiplied by 0.2 to derive the WFC).

Users of natureindex.com can search for specific institutions or countries and generate their own reports, ordered by article count (AC), fractional count (FC) or weighted fractional count (WFC).

Each query will return a profile page that lists the country or institution's recent research outputs, from which it is possible to drill down for more information. For example, articles can be displayed by journal, and then by article title. As in the supplement, research outputs are organized by subject area. The profile page also lists the institution or country's top collaborators, as well as its relationship with other research organizations.

The total FC or WFC for an institution is derived by summing the FC or WFC for individual authors. The process is similar for countries, although complicated by the fact that some institutions have overseas labs that will be counted towards the host country totals. What's more, there is great variability in the way authors present their affiliations. Every effort is made to count affiliations consistently, making reasonable assumptions. For more information on how the affiliation information is processed, please see the frequently asked questions at natureindex.com.

THE SUPPLEMENT

Nature Index 2014 China is based on a snapshot of data from natureindex.com, covering articles published between 1 January and 31 December, 2013.

Most analyses within the Nature Index 2014

China supplement use the WFC as the primary metric, as it provides a more even basis for comparison across multiple disciplines, and in determining the relative contribution of each country/institution.

Additional layers of information concerning funding levels, numbers of researchers, size of population and so on, are taken from publicly available sources. In several places, we use altmetrics as a supporting data source. Altmetrics is an alternative way to measure the impact of a paper by tracking different online sources (newspaper stories, tweets, blog posts, comments) that mention the paper. The altmetric score for an article gives an idea of the attention that it has received. Our data are from altmetric.com, provided by the start-up company Altmetric — which is supported by Digital Science. To see more about how this score is calculated, please visit support.altmetric.com. ■

NATUREINDEX.COM

A global indicator of high-quality research

Institution name

Country

Research

Collaboration

Relationships

September 2013 - August 2014

Region: Global
Subject/journal group: All

The table to the right includes counts of all research outputs for Institution name published between September 2013 – August 2014 which are tracked by the Nature Index.

Below, the same research outputs are grouped by subject. Click on the subject to drill-down into a list of articles organized by journal, and then by title.

Note: Articles may be assigned to more than one subject area.

AC	FC	WFC
1221	598.04	558.30

Outputs by subject



Subject	AC	FC	WFC
Chemistry	276	179.1	179.11
Earth & Environmental Sciences	95	42.73	42.73
Life Sciences	439	231.50	231.50
Physical Sciences	652	284.48	244.74

Return to institution outputs

Nature Index China tables

China's leading institutions for high-quality science, ordered by weighted fractional count (WFC) for 2013. Also shown are the total number of articles, and the change in WFC from 2012. Articles are from the 68 natural science journals that comprise the Nature Index (see 'A guide to the Nature Index', page S76).

TOP 200 INSTITUTIONS

2013	INSTITUTION	WFC	ARTICLE COUNT	2012 WFC	2012-2013 CHANGE IN WFC
1	Chinese Academy of Sciences (CAS)	1,209.46	2,661	1,119.75	8.0%
2	Peking University (PKU)	275.53	743	209.58	31.5%
3	Tsinghua University	194.87	474	177.74	9.6%
4	Nanjing University	194.57	391	168.10	15.7%
5	University of Science and Technology of China (USTC)	175.73	427	147.75	18.9%
6	Zhejiang University (ZJU)	150.42	289	123.20	22.1%
7	Fudan University	129.23	255	121.36	6.5%
8	Nankai University	113.77	190	85.85	32.5%
9	Wuhan University	98.80	154	74.27	33.0%
10	Jilin University	97.90	179	65.76	48.9%
11	Shanghai Jiao Tong University (SJTU)	95.99	247	80.03	19.9%
12	Sun Yat-sen University	79.41	158	80.04	-0.8%
13	Sichuan University	76.82	130	44.88	71.2%
14	Xiamen University	76.02	142	77.84	-2.3%
15	University of Chinese Academy of Sciences (UCAS)	71.18	434	64.84	9.8%
16	The University of Hong Kong (HKU)	70.43	149	50.45	39.6%
17	Lanzhou University	69.99	123	67.58	3.6%
18	East China Normal University (ECNU)	65.56	123	35.55	84.4%
19	Soochow University	65.30	128	55.27	18.1%
20	Dalian University of Technology (DUT)	61.42	104	51.85	18.5%
21	East China University of Science and Technology (ECUST)	56.75	95	67.33	-15.7%
22	Hunan University (HNU)	54.57	80	50.22	8.7%
23	Hong Kong University of Science and Technology (HKUST)	54.45	103	55.07	-1.1%
24	Huazhong University of Science & Technology (HUST)	43.62	109	46.44	-6.1%
25	Xi'an Jiaotong University	42.98	93	30.27	42.0%
26	Tongji University	40.08	86	20.29	97.5%
27	Beijing Normal University	39.60	121	36.73	7.8%
28	The Chinese University of Hong Kong	39.39	82	40.20	-2.0%
29	Shandong University	39.17	94	63.60	-38.4%
30	City University of Hong Kong	36.51	72	37.02	-1.4%
31	Harbin Institute of Technology	36.22	73	19.30	87.6%
32	Tianjin University	33.73	66	38.57	-12.6%
33	Southeast University (SEU)	30.93	65	25.77	20.0%
34	South China University of Technology	30.89	60	32.05	-3.6%
35	Northeast Normal University	30.73	48	35.43	-13.3%
36	Fuzhou University	26.76	39	29.64	-9.7%
37	The Hong Kong Polytechnic University	25.94	69	26.09	-0.6%
38	University of Science and Technology Beijing (USTB)	25.79	46	16.71	54.3%
39	Chinese Academy of Medical Sciences & Peking Union Medical College (CAMS & PUMC)	24.63	75	23.32	5.6%
40	Beijing University of Chemical Technology	23.45	40	16.63	41.0%
41	Beijing Institute of Technology	20.11	40	15.20	32.3%
42	Beihang University (BUAA)	17.69	67	18.59	-4.8%

2013	INSTITUTION	WFC	ARTICLE COUNT	2012 WFC	2012-2013 CHANGE IN WFC
43	Second Military Medical University	16.96	44	10.84	56.4%
44	Shanghai University	16.49	41	21.87	-24.6%
45	National University of Defense Technology	16.24	29	12.00	35.3%
46	Southwest University	16.11	33	18.37	-12.3%
47	Central China Normal University	15.93	43	25.10	-36.5%
48	BGI	15.34	51	10.56	45.2%
49	China Agricultural University	14.65	41	8.49	72.6%
50	China University of Geosciences	14.52	40	13.48	7.7%
51	Zhengzhou University	13.77	38	7.78	77.1%
52	Ocean University of China	13.06	35	13.20	-1.1%
53	Northwestern Polytechnical University (NWPUP)	12.90	18	1.82	610.4%
54	Chongqing University	12.90	22	3.00	329.9%
55	Xiangtan University	12.87	27	12.92	-0.4%
56	Hong Kong Baptist University (HKBU)	12.77	30	10.60	20.4%
57	Shandong Normal University (SDNU)	12.73	18	8.83	44.2%
58	Henan University	12.53	18	8.14	53.9%
59	Henan Normal University	12.42	27	9.95	24.9%
60	Northwest University	12.04	25	11.48	4.9%
61	University of Electronic Science and Technology of China (UESTC)	12.00	32	12.42	-3.3%
62	Nanjing University of Technology	11.92	28	16.16	-26.3%
63	Shaanxi Normal University	11.80	24	3.51	236.5%
64	National Institute of Biological Sciences (NIBS)	11.73	30	15.84	-26.0%
65	Yunnan University	11.39	27	7.36	54.9%
66	Hunan Normal University	11.12	24	10.10	10.1%
67	Shanxi University	10.82	18	16.00	-32.4%
68	Qingdao University of Science and Technology (QUST)	10.66	17	14.16	-24.7%
69	Huazhong Agricultural University	10.57	20	8.38	26.2%
70	Central South University (CSU)	10.15	35	6.46	57.3%
71	Nanchang University	10.06	20	3.69	172.3%
72	Northwest A & F University	9.90	21	3.57	177.1%
73	Hangzhou Normal University	9.52	36	6.55	45.3%
74	China Earthquake Administration	9.50	20	4.43	114.4%
75	Hefei University of Technology	8.97	18	3.48	157.9%
76	Jiangnan University	8.91	17	8.03	10.9%
77	Renmin University of China	8.89	20	10.51	-15.4%
78	Wuhan University of Technology	8.59	16	7.78	10.4%
79	Zhejiang Normal University (ZJNU)	8.44	21	6.04	39.9%
80	Nanjing University of Posts and Telecommunications (NUPT)	8.28	18	3.00	176.1%
81	Wenzhou University	8.05	15	5.17	55.7%
82	Zhejiang University of Technology	8.03	15	6.50	23.5%
83	Chinese Academy of Agricultural Sciences (CAAS)	7.87	30	4.90	60.7%
84	Hebei University	7.86	12	3.42	129.9%
85	Nanjing University of Information Science & Technology (NUIST)	7.86	30	6.67	17.9%
86	Nanjing Normal University	7.75	23	9.95	-22.1%
87	Beijing University of Technology	7.67	16	7.45	3.0%
88	Ningbo University	7.56	20	3.52	114.9%
89	China Academy of Engineering Physics (CAEP)	7.32	21	5.79	26.4%
90	Nanjing Medical University	7.27	28	7.97	-8.7%
91	China University of Petroleum (CUP)	7.15	16	5.25	36.0%
92	Shantou University	7.13	14	5.76	23.7%
93	South China Normal University	6.87	18	12.13	-43.4%
94	Tianjin Medical University (TMC)	6.73	26	5.70	18.1%
95	China Meteorological Administration (CMA)	6.69	26	4.50	48.6%
96	Heilongjiang University	6.67	14	5.46	22.1%
97	Capital Normal University	6.65	16	5.18	28.6%
98	Nanjing University of Aeronautics and Astronautics (NUAA)	6.58	10	15.17	-56.6%
99	North China Electric Power University (NCEPU)	6.50	13	2.16	201.3%
100	Nanjing University of Science and Technology (NUST)	6.43	18	7.77	-17.2%

2013	INSTITUTION	WFC	ARTICLE COUNT	2012 WFC	2012-2013 CHANGE IN WFC
101	Donghua University	6.34	18	1.78	257.1%
102	Anhui Normal University	5.96	9	8.81	-32.4%
103	Beijing Computational Science Research Center (CSRC)	5.95	24	2.37	151.6%
104	State Oceanic Administration (SOA)	5.93	20	9.74	-39.1%
105	Changzhou University	5.93	13	0.56	967.3%
106	Third Military Medical University	5.91	25	3.81	55.0%
107	Jiangsu Normal University	5.61	11	2.17	158.8%
108	University of Jinan	5.57	6	4.58	21.5%
109	Tianjin University of Technology (TUT)	5.53	13	1.06	420.3%
110	Huaqiao University	5.51	11	2.48	121.7%
111	Hunan University of Science and Technology	5.06	8	-	-
112	Shenzhen University (SZU)	5.05	15	2.43	108.0%
113	Chinese Academy of Geological Sciences (CAGS)	5.00	16	4.11	21.6%
114	Zhejiang Sci-Tech University	4.94	9	3.37	46.9%
115	Huaibei Normal University	4.78	7	5.20	-8.2%
116	Yangzhou University	4.67	9	3.94	18.4%
117	Guangxi Normal University	4.60	12	10.13	-54.6%
118	Taiyuan University of Technology (TUT)	4.54	8	1.38	230.0%
119	Jinan University	4.31	21	9.90	-56.5%
120	Chongqing Medical University (CQMU)	4.31	14	2.60	66.0%
121	University of Shanghai for Science and Technology (USST)	4.15	11	2.25	84.6%
122	Beijing Jiaotong University	4.15	7	1.81	129.2%
123	Yanshan University	4.03	7	6.62	-39.2%
124	China University of Mining and Technology (CUMT)	3.95	8	3.19	24.0%
125	China Pharmaceutical University (CPU)	3.88	10	8.66	-55.2%
126	Fourth Military Medical University	3.80	14	5.17	-26.4%
127	Xidian University	3.78	8	3.52	7.5%
128	Shanghai Normal University (SHNU)	3.75	20	9.11	-58.8%
129	Anhui University	3.71	11	1.83	102.9%
130	Academy of Military Medical Sciences (AMMS)	3.67	13	4.74	-22.4%
131	Qufu Normal University (QFNU)	3.63	7	-	-
132	Hubei University	3.63	10	1.83	98.0%
133	Inner Mongolia University	3.61	9	0.74	386.1%
134	Fujian Normal University	3.48	11	3.84	-9.3%
135	Southern Medical University	3.47	12	1.78	94.3%
136	Jiangsu University	3.42	13	3.39	0.9%
137	Beijing University of Posts and Telecommunications (BUPT)	3.42	8	4.11	-16.9%
138	China Medical University (PRC)	3.32	16	1.08	207.0%
139	Linyi University	3.19	6	0.63	411.1%
140	Guizhou University	3.18	8	1.48	115.4%
141	Jiangxi Normal University	3.14	6	3.58	-12.3%
142	Hohai University	3.04	9	0.84	263.5%
143	Kunming University of Science and Technology	3.04	7	0.06	5,214.6%
144	Institute of Applied Physics and Computational Mathematics (IAPCM)	3.00	11	5.20	-42.3%
145	PLA University of Science and Technology	3.00	4	1.43	110.5%
146	Nantong University	2.99	9	2.12	41.2%
147	Capital Medical University (CMU)	2.98	16	2.28	30.5%
148	Henan University of Technology (HUT)	2.92	6	1.10	165.6%
149	Chinese Center for Disease Control and Prevention (China CDC)	2.91	17	2.46	18.4%
150	Qingdao University	2.89	8	2.05	40.9%

2013	INSTITUTION	WFC	ARTICLE COUNT	2012 WFC	2012-2013 CHANGE IN WFC
151	Beijing Institute of Biotechnology	2.81	8	1.40	100.9%
152	Sichuan Agricultural University (SAU)	2.70	5	0.65	316.3%
153	Tianjin University of Science and Technology (TUST)	2.69	6	1.11	143.0%
154	Nanjing Agricultural University	2.59	11	3.69	-29.7%
155	Harbin Medical University (HMU)	2.58	11	1.73	48.7%
156	Chinese Academy of Meteorological Sciences (CAMS)	2.53	9	5.06	-50.0%
157	Harbin Engineering University (HEU)	2.52	5	0.44	467.3%
158	China Jiliang University	2.38	8	2.03	17.1%
159	Henan University of Science and Technology	2.37	9	0.43	454.2%
160	Liaocheng University (LCU)	2.29	6	0.80	185.0%
161	South University of Science and Technology of China (SUSTC)	2.24	22	0.62	260.8%
162	Nanjing National Laboratory of Microstructures	2.18	5	2.41	-9.5%
163	Xi'an University of Technology (XUT)	2.15	4	1.10	96.7%
164	Beijing Institute of Pharmacology and Toxicology	2.13	4	1.75	21.5%
165	Changchun University of Science and Technology (CUST)	2.12	8	2.63	-19.4%
166	Northeastern University	2.10	4	6.25	-66.4%
167	South China Agricultural University	2.02	6	1.68	20.6%
168	Guangzhou Medical University (GMU)	2.01	7	1.25	61.4%
169	Shanxi Datong University (SDU)	2.00	2	-	-
170	Dalian National Laboratory for Clean Energy (DNL)	1.99	4	-	-
171	Inner Mongolia University of Science and Technology	1.95	3	0.33	485.0%
172	Yunnan University of Nationalities	1.93	3	0.78	147.5%
173	Yunnan Normal University (YNNU)	1.91	11	0.79	143.5%
174	JiangXi University of Science and Technology (JUST)	1.91	4	-	-
175	China West Normal University (CWNU)	1.82	3	3.29	-44.7%
176	Shanghai Second Polytechnic University	1.81	3	-	-
177	Hebei Semiconductor Research Institute	1.77	2	0.30	489.7%
178	Anhui Medical University	1.77	13	1.92	-8.1%
179	Shenyang Normal University	1.75	6	0.64	172.2%
180	Nanchang Hangkong University (NCHU)	1.73	4	1.18	46.1%
181	Gannan Normal University	1.71	2	0.63	174.3%
182	Shanghai Institute of Technology	1.71	4	-	-
183	Fujian Medical University	1.71	4	1.08	57.7%
184	Guangzhou University (GU)	1.70	9	0.36	369.1%
185	Qingdao Agricultural University	1.68	5	-	-
186	Harbin University of Science and Technology (HUST)	1.67	5	-	-
187	Hebei University of Technology (HEBUT)	1.65	8	0.34	378.6%
188	Xinjiang University	1.59	11	2.61	-39.1%
189	Guilin University of Electronic Technology (GUET)	1.56	6	1.54	1.2%
190	Changshu Institute of Technology (CIT)	1.55	5	1.51	2.6%
191	Wuyi University	1.54	2	0.68	125.9%
192	The General Hospital of Chinese People's Liberation Army	1.54	11	1.58	-2.8%
193	Wenzhou Medical College	1.49	9	3.81	-60.9%
194	Guangdong University of Technology (GDUT)	1.48	7	0.36	307.9%
195	Guilin University of Technology	1.48	6	3.14	-53.0%
196	Guangdong Medical College	1.45	5	0.76	90.2%
197	Jiangxi Science & Technology Normal University	1.43	3	1.25	14.0%
198	Changchun University of Technology (CCUT)	1.42	4	0.40	255.7%
199	Hebei Normal University	1.40	12	1.86	-24.6%
200	Shandong Agricultural University (SDAU)	1.38	6	2.00	-31.3%

TOP INSTITUTIONS: LIFE SCIENCES

2013	INSTITUTION	WFC	ARTICLE COUNT	2012 WFC	2012-2013 CHANGE IN WFC
1	Chinese Academy of Sciences (CAS)	238.91	478	201.26	18.7%
2	Peking University (PKU)	64.01	152	42.93	49.1%
3	Tsinghua University	30.65	87	34.02	-9.9%
4	University of Science and Technology of China (USTC)	30.21	62	14.98	101.7%
5	Zhejiang University (ZJU)	29.29	68	27.26	7.5%
6	Nankai University	27.74	54	20.83	33.2%
7	Shanghai Jiao Tong University (SJTU)	24.82	94	24.08	3.1%
8	Fudan University	24.72	68	29.10	-15.1%
9	Wuhan University	22.07	40	16.80	31.4%
10	Chinese Academy of Medical Sciences & Peking Union Medical College (CAMS & PUMC)	20.51	66	22.34	-8.2%
11	Lanzhou University	19.26	33	25.75	-25.2%
12	Sun Yat-sen University	17.74	55	18.81	-5.7%
13	East China Normal University (ECNU)	16.26	34	5.62	189.0%
14	Sichuan University	16.14	34	8.21	96.6%
15	BGI	14.82	49	10.09	46.9%
16	Second Military Medical University	14.77	38	9.38	57.5%
17	University of Chinese Academy of Sciences (UCAS)	14.46	84	9.84	46.9%
18	Soochow University	13.24	30	9.50	39.4%
19	Nanjing University	12.67	37	11.00	15.1%
20	The University of Hong Kong (HKU)	12.03	31	17.81	-32.4%
21	Tongji University	10.76	32	3.90	175.9%
22	Huazhong University of Science & Technology (HUST)	10.51	30	10.16	3.4%
23	Xiamen University	10.26	25	14.95	-31.4%
24	National Institute of Biological Sciences (NIBS)	9.77	26	15.22	-35.8%
25	East China University of Science and Technology (ECUST)	9.67	17	9.98	-3.1%
26	South China University of Technology	9.33	15	5.66	64.9%
27	China Agricultural University	9.16	28	6.06	51.1%
28	Northeast Normal University	8.99	17	4.06	121.4%
29	Shandong University	8.90	18	17.22	-48.3%
30	Hong Kong University of Science and Technology (HKUST)	8.48	15	14.31	-40.7%
31	Huazhong Agricultural University	7.61	15	5.32	43.1%
32	Dalian University of Technology (DUT)	6.51	14	5.18	25.7%
33	Chinese Academy of Agricultural Sciences (CAAS)	6.39	24	4.63	37.9%
34	Wenzhou University	5.90	8	2.61	126.3%
35	Nanjing Medical University	5.86	23	5.71	2.7%
36	Tianjin Medical University (TMC)	5.79	21	4.95	17.0%
37	Beijing Normal University	5.79	21	8.80	-34.2%
38	The Chinese University of Hong Kong	5.60	19	10.89	-48.6%
39	Tianjin University	5.06	8	3.59	40.8%
40	Central China Normal University	4.16	5	6.00	-30.7%
41	Third Military Medical University	4.12	19	2.81	46.6%
42	Jiangsu Normal University	4.03	5	0.75	437.8%
43	Zhengzhou University	3.89	11	2.42	60.7%
44	The Hong Kong Polytechnic University	3.89	14	2.74	41.7%
45	Northwest A & F University	3.70	12	1.83	102.9%
46	Fourth Military Medical University	3.42	12	4.17	-18.0%
47	Yunnan University	3.11	7	2.78	12.1%
48	Southern Medical University	3.10	10	1.45	113.6%
49	Huaqiao University	3.08	4	0.08	3,600.0%
50	Henan University	3.02	5	1.59	89.9%

TOP INSTITUTIONS: CHEMISTRY

2013	INSTITUTION	WFC	ARTICLE COUNT	2012 WFC	2012-2013 CHANGE IN WFC
1	Chinese Academy of Sciences (CAS)	679.85	1,148	654.95	3.8%
2	Peking University (PKU)	142.80	268	113.77	25.5%
3	Nanjing University	115.65	160	97.50	18.6%
4	University of Science and Technology of China (USTC)	93.78	150	82.17	14.1%
5	Tsinghua University	92.00	169	90.79	1.3%
6	Zhejiang University (ZJU)	86.42	133	69.32	24.7%
7	Nankai University	86.25	124	59.50	44.9%
8	Fudan University	79.94	122	71.59	11.7%
9	Jilin University	74.06	129	40.87	81.2%
10	Sichuan University	68.45	92	33.30	105.5%
11	Wuhan University	63.22	84	42.23	49.7%
12	Xiamen University	59.87	95	58.89	1.7%
13	East China University of Science and Technology (ECUST)	53.04	84	64.81	-18.2%
14	Sun Yat-sen University	52.10	75	42.79	21.8%
15	Hunan University (HNU)	51.90	68	43.66	18.9%
16	Lanzhou University	50.35	74	47.97	5.0%
17	University of Chinese Academy of Sciences (UCAS)	47.85	229	42.67	12.1%
18	Dalian University of Technology (DUT)	47.16	75	42.72	10.4%
19	Shanghai Jiao Tong University (SJTU)	45.14	72	35.41	27.5%
20	East China Normal University (ECNU)	41.54	71	20.52	102.4%
21	The University of Hong Kong (HKU)	38.04	52	21.38	77.9%
22	Soochow University	34.75	70	37.77	-8.0%
23	South China University of Technology	28.89	51	26.26	10.0%
24	Fuzhou University	26.51	37	28.48	-6.9%
25	Hong Kong University of Science and Technology (HKUST)	25.71	44	29.59	-13.1%
26	Northeast Normal University	25.69	37	31.95	-19.6%
27	Tianjin University	25.43	47	26.03	-2.3%
28	The Chinese University of Hong Kong	21.18	34	22.56	-6.1%
29	Beijing University of Chemical Technology	20.35	31	14.71	38.3%
30	Shandong University	19.69	32	24.70	-20.3%
31	Tongji University	19.01	32	9.74	95.3%
32	City University of Hong Kong	18.84	36	15.60	20.8%
33	Beijing Institute of Technology	16.95	28	10.87	55.9%
34	The Hong Kong Polytechnic University	15.16	37	21.26	-28.7%
35	Harbin Institute of Technology	14.15	34	10.18	39.1%
36	Central China Normal University	13.30	18	21.92	-39.3%
37	Huazhong University of Science & Technology (HUST)	13.11	27	18.82	-30.4%
38	Southwest University	12.51	23	17.92	-30.2%
39	Shandong Normal University (SDNU)	11.33	13	7.73	46.6%
40	Zhengzhou University	11.29	20	6.61	70.9%
41	Chinese Academy of Medical Sciences & Peking Union Medical College (CAMS & PUMC)	11.22	22	7.36	52.5%
42	Nanjing University of Technology	11.04	24	15.67	-29.5%
43	Henan University	10.73	14	4.69	128.9%
44	Xi'an Jiaotong University	10.72	26	6.97	53.8%
45	Qingdao University of Science and Technology (QUST)	10.66	17	14.05	-24.1%
46	Beijing Normal University	10.56	23	12.71	-16.9%
47	Northwest University	10.29	17	7.18	43.3%
48	Southeast University (SEU)	9.78	21	11.15	-12.3%
49	Shaanxi Normal University	9.14	17	2.36	286.7%
50	Beihang University (BUAA)	8.60	22	8.94	-3.8%

TOP INSTITUTIONS: PHYSICAL SCIENCES

2013	INSTITUTION	WFC	ARTICLE COUNT	2012 WFC	2012-2013 CHANGE IN WFC
1	Chinese Academy of Sciences (CAS)	412.33	1,234	399.92	3.1%
2	Peking University (PKU)	105.35	388	86.35	22.0%
3	Tsinghua University	88.12	262	87.16	1.1%
4	University of Science and Technology of China (USTC)	71.66	242	63.63	12.6%
5	Zhejiang University (ZJU)	65.02	136	44.46	46.2%
6	Nanjing University	64.94	193	68.24	-4.8%
7	Fudan University	40.00	92	41.84	-4.4%
8	Shanghai Jiao Tong University (SJTU)	34.39	104	29.26	17.5%
9	Xi'an Jiaotong University	32.74	56	24.34	34.5%
10	Jilin University	29.09	54	25.59	13.7%
11	Soochow University	27.31	45	22.45	21.6%
12	The University of Hong Kong (HKU)	25.81	70	15.56	65.9%
13	Huazhong University of Science & Technology (HUST)	25.28	61	26.04	-2.9%
14	Nankai University	23.56	46	17.53	34.4%
15	Harbin Institute of Technology	22.37	40	9.59	133.3%
16	Hong Kong University of Science and Technology (HKUST)	22.11	49	18.87	17.2%
17	Sun Yat-sen University	21.84	40	21.84	0.0%
18	University of Science and Technology Beijing (USTB)	21.29	38	11.55	84.4%
19	City University of Hong Kong	20.40	44	24.73	-17.5%
20	Southeast University (SEU)	20.02	42	14.59	37.2%
21	Beijing Normal University	19.47	63	15.69	24.1%
22	Dalian University of Technology (DUT)	18.98	35	8.83	114.8%
23	Wuhan University	18.52	37	17.49	5.9%
24	East China Normal University (ECNU)	17.17	37	13.07	31.3%
25	University of Chinese Academy of Sciences (UCAS)	16.51	158	17.51	-5.7%
26	National University of Defense Technology	16.16	28	11.98	34.8%
27	Shandong University	14.13	49	28.93	-51.2%
28	Tongji University	13.22	25	7.70	71.7%
29	Lanzhou University	13.15	30	14.36	-8.4%
30	The Chinese University of Hong Kong	13.12	30	11.06	18.6%
31	Xiamen University	12.86	35	11.09	15.9%
32	Northwestern Polytechnical University (NWPU)	11.74	16	1.27	826.8%
33	University of Electronic Science and Technology of China (UESTC)	11.41	26	12.16	-6.2%
34	Beihang University (BUAA)	10.08	50	12.89	-21.8%
35	Tianjin University	9.98	23	14.49	-31.1%
36	The Hong Kong Polytechnic University	9.80	26	9.24	6.0%
37	Chongqing University	9.50	17	2.00	374.8%
38	Shanghai University	8.28	25	9.82	-15.7%
39	Henan Normal University	7.91	21	3.60	119.9%
40	Beijing Institute of Technology	7.84	21	7.10	10.5%
41	Sichuan University	7.74	25	9.90	-21.8%
42	East China University of Science and Technology (ECUST)	6.99	15	1.61	333.7%
43	Shanxi University	6.06	12	8.06	-24.8%
44	Wuhan University of Technology	5.78	11	4.09	41.3%
45	North China Electric Power University (NCEPU)	5.54	10	2.09	164.7%
46	Beijing University of Technology	5.46	10	6.10	-10.5%
47	Xiangtan University	5.39	14	5.76	-6.5%
48	China Academy of Engineering Physics (CAEP)	5.36	14	4.89	9.6%
49	Beijing Computational Science Research Center (CSRC)	5.13	20	2.03	152.5%
50	Ningbo University	5.07	14	3.52	44.2%

TOP INSTITUTIONS: EARTH AND ENVIRONMENTAL SCIENCES

2013	INSTITUTION	WFC	ARTICLE COUNT	2012 WFC	2012-2013 CHANGE IN WFC
1	Chinese Academy of Sciences (CAS)	68.50	133	44.78	53.0%
2	China Earthquake Administration	9.50	20	4.27	122.3%
3	China University of Geosciences	9.40	20	7.80	20.6%
4	Ocean University of China	9.35	23	7.92	18.1%
5	Nanjing University	9.33	18	2.31	304.2%
6	China Meteorological Administration	6.52	23	4.32	51.1%
7	University of Science and Technology of China (USTC)	6.52	15	2.32	180.5%
8	Nanjing University of Information Science & Technology (NUIST)	6.30	22	6.59	-4.5%
9	Wuhan University	5.93	12	3.93	50.8%
10	Beijing Normal University	5.76	19	3.70	55.7%
11	Peking University (PKU)	5.61	16	6.00	-6.5%
12	State Oceanic Administration (SOA)	5.18	14	9.74	-46.9%
13	Lanzhou University	5.18	11	3.21	61.4%
14	Hong Kong University of Science and Technology (HKUST)	4.54	8	2.29	98.2%
15	Chinese Academy of Geological Sciences (CAGS)	3.77	12	3.96	-4.6%
16	University of Chinese Academy of Sciences (UCAS)	2.96	20	3.75	-21.1%
17	PLA University of Science and Technology	2.80	3	1.43	96.5%
18	Chinese Academy of Meteorological Sciences (CAMS)	2.53	9	4.89	-48.3%
19	Tsinghua University	2.46	5	0.11	2,083.4%
20	Sun Yat-sen University	2.29	10	2.90	-21.3%
21	The Hong Kong Polytechnic University	2.25	5	0.96	134.2%
22	Zhejiang University (ZJU)	1.88	5	1.63	15.2%
23	Xiamen University	1.80	6	2.21	-18.6%
24	Hohai University	1.44	5	0.84	71.6%
25	The Chinese University of Hong Kong	1.15	4	2.33	-50.9%

TOP INSTITUTIONS IN NATURE AND SCIENCE

2013	INSTITUTION	WFC	ARTICLE COUNT	2012 WFC	2012-2013 CHANGE IN WFC
1	Chinese Academy of Sciences (CAS)	18.64	54	6.26	197.5%
2	Tsinghua University	5.43	12	5.56	-2.2%
3	Peking University (PKU)	4.10	14	3.56	15.2%
4	Chinese Academy of Agricultural Sciences (CAAS)	2.84	7	0.74	282.8%
5	BGI	1.78	7	1.33	33.6%
6	Zhejiang University (ZJU)	1.70	7	0.53	221.1%
7	University of Science and Technology of China (USTC)	1.69	8	2.13	-20.8%
8	Chinese Center for Disease Control and Prevention (China CDC)	1.27	6	-	-
9	China Agricultural University	1.17	6	0.32	269.6%
10	National Institute of Biological Sciences, (NIBS)	1.12	5	2.60	-56.8%
11	Tongji University	1.11	5	0.06	1,717.6%
12	Linyi University	1.05	3	-	-
13	Chinese Academy of Geological Sciences (CAGS)	0.98	2	0.10	876.9%
14	The University of Hong Kong (HKU)	0.82	5	0.23	263.4%
15	Yunnan University	0.80	2	0.38	113.3%
16	Fudan University	0.76	3	0.42	81.8%
17	Yanshan University	0.73	1	-	-
18	Shantou University	0.67	2	-	-
19	Shenyang Normal University	0.58	2	0.17	250.0%
20	Dalian University of Technology (DUT)	0.55	3	0.02	2,326.2%
21	Southeast University (SEU)	0.55	1	0.29	90.9%
22	University of Chinese Academy of Sciences (UCAS)	0.50	5	0.41	24.4%
23	Shandong Tianyu Natural History Museum	0.49	3	-	-
24	China University of Petroleum (CUP)	0.48	1	-	-
25	China Medical University (PRC)	0.43	1	0.13	242.9%

Weighted fractional count (WFC) for each institution is shown to two decimal places only. When two or more institutions have the same WFC, their positions are determined by the thousandth place (or beyond).

These results are based on the most recent data available as of 11 September 2014. Owing to continual refinements of the data, the figures in the database are liable to change and might differ to those printed in the supplements.



CAS-MPG Partner Institute for Computational Biology

Establishment

The CAS-MPG Partner Institute for Computational Biology (PICB) is a research institute located in the heart of Shanghai. PICB was jointly established in 2005 by the Max Planck Society (MPG) of Germany and the Chinese Academy of Sciences (CAS), and is operated under their joint guidance. PICB is dedicated to broad topics in the quantitative biosciences with a particular focus on computational biology. The institute's purpose is to explore scientific frontiers, contribute to the education and training of excellent junior scientists, and complement scientific research conducted at the research institutes of CAS and the MPG.

Like the Max Planck Institutes, PICB is headed by a board of directors comprising directors and department heads and, as of recently, other institute representatives. The difference in this organizational model to a typical CAS institute is that there are several directors rather than one institute director. Directors are selected by the CAS-MPG Joint Core Commission and are appointed by CAS. PICB enjoys complete autonomy with regard to scientific focus; however, its research work is subject to continuous supervision and evaluation by a scientific advisory board.

Structure of PICB

PICB has a relatively flat structure consisting of the following units:

- departments headed by directors
- research groups headed by principal investigators (PIs)
- Max Planck Independent Research Groups (IRGs)

A department is typically larger than an IRG or a PI group (which stand by themselves). Directors and IRG heads are selected by the CAS-MPG Joint Core Commission and have their budgets directly allocated by CAS and the MPG. PIs are recruited according to strict scientific standards following a CAS procedure, which relies on the vote of a committee of CAS PIs, consisting of both PICB directors and PIs, and at least a third of PIs external to PICB. PICB is managed by the board of directors, which consists of PICB directors, a representative of the IRG heads and the PIs, plus a representative of the Shanghai Institutes for Biological Sciences (SIBS), which is a collection of eight local CAS institutes focusing on biological research.

Research Concept

PICB merges theoretical and experimental biology in three focal areas:

1. Integrative analysis of gene regulation
2. Computational modeling of complex traits
3. Computational analysis of human variation and evolution

In more theoretical terms, the research can be summarized as **translating large-scale multi-omics data into novel knowledge of human biology**. The 'multi-omics data' includes diverse data-types such as transcriptome data, gene regulation data, epigenetic data, proteomics, and metabolomics data. Computational biology today largely deals with these data-types, their analysis, integration and interpretation. Therefore, **PICB strives to seamlessly integrate computational and experimental biology to understand biological processes through quantitative approaches**.

THE CAS-MPG PARTNER INSTITUTE FOR COMPUTATIONAL BIOLOGY IS HIRING 'GROUP LEADERS' IN SHANGHAI, CHINA

The CAS-MPG Partner Institute for Computational Biology is an internationally recognized research institute based in Shanghai, China — jointly operated by the Chinese Academy of Sciences (CAS) and the German Max Planck Society (MPG). Work at this institute is driven by the growing importance of statistical and computational methods in modern biology. We undertake innovative research in the interdisciplinary fields of biology, mathematics, physics and computer science.

PICB is eager to receive applications for several group leader positions from talented scientists working in the broad field of quantitative biology, including but not limited to the following areas:

- genome biology
- epigenomics and RNA processing
- computational biology
- biostatistics
- biomathematics

PICB will offer a competitive salary package to successful applicants, including a basic salary, position allowance, housing allowance and other benefits. There is no deadline and applicant evaluations will remain open until all positions are filled.

Interested applicants should send a covering letter, curriculum vitae and brief summary of past research achievements and future plans, accompanied by three letters of recommendation to:

Professor Jing-Dong Jackie Han
CAS-MPG Partner Institute for Computational Biology
320 Yueyang Road
Shanghai 200031
Phone: 86-21-54920458
Fax: 86-21-54920451
E-mail: jdhan@picb.ac.cn





THE GUANGZHOU INSTITUTES OF BIOMEDICINE AND HEALTH

A CENTRE FOR TRANSLATIONAL MEDICINE IN THE PEARL RIVER DELTA

The Guangzhou Institutes of Biomedicine and Health (GIBH) was established in 2003 by the Chinese Academy of Sciences through collaborative agreements with the Guangdong and Guangzhou governments. It was envisioned to be a modern institute with a forward-looking philosophy that would transform the research landscape in southern China. We are very proud to report that we are realizing this vision of our founders.

We have built a strong organization with approximately 500 dedicated staff members and more than 280 graduate students. Our scientists and students are engaged in three major areas of inquiry — stem cell biology, chemical and synthetic biology, and infection and immunity. We believe that basic research in these three areas will not only enhance our understanding of life and disease,

but also catalyse breakthroughs in disease diagnosis and treatment. Based on this belief, we are devoting our resources to developing cutting-edge technologies and solving important scientific problems in a collaborative environment. Accordingly, we have established a centralized platform for drug discovery and technology development. We are also forming a new unit on public health, which will help us learn more about the impact of social-economic developments on human health in southern China.

Our research has led to high-impact publications in leading academic journals. Our innovative drug-discovery platform, in particular, has advanced several innovative concepts for treating cancer and neural degenerative diseases and has developed drug candidates that are poised for clinical testing and registration. Our dual focus on original

discovery and innovative applications will drive our organization towards even greater achievements in the coming years.

Our mission at GIBH is to serve the citizens of China and the world through scientific discovery. We are looking forward to forming external collaborations and partnerships with scientists and institutes to work towards achieving common goals. We are proud that GIBH represents commitment, creativity and opportunity.

We seek individuals who share our vision and enthusiasm for the future. Positions are available at all levels in our five research programmes:

- Stem cell and regenerative medicine
- Chemistry and synthetic biology
- Infection and immunity
- Public health
- Drug discovery pipeline

For more information, please visit us at
<http://www.gibh.cas.cn> or
<http://english.gibh.cas.cn>



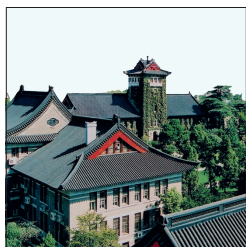
First graduates from GIBH celebrating the successful completion of their studies.



LEARN MORE

Visit: <http://english.gibh.cas.cn>
 Email: hr@gibh.org
 Address: 190 Kaiyuan Avenue,
 Science Park, Guangzhou,
 China, 510530

Tel: +86-20-32015342
 Fax: +86-20-32015267



:insideview

profile feature



Please visit our website at www.microstructures.org
or contact us at wudi@nju.edu.cn

The Collaborative Innovation Center of Advanced Microstructures (CICAM) was formally authenticated by the Ministry of Education of China in 2014. Here, we discuss CICAM's mission and development with three CICAM directors: Dingyu Xing of Nanjing University, Fuchun Zhang of Zhejiang University and Xingao Gong of Fudan University.

Q: Who founded the centre?

Nanjing University took the lead in founding CICAM in 2012, in partnership with Fudan University and Shanghai Jiao Tong University (both in Shanghai), Zhejiang University (in Hangzhou), the University of Science and Technology of China and the Hefei Institutes of Physical Science of the Chinese Academy of Sciences (both in Hefei) and the company Huawei Technologies. The cities of Shanghai, Hefei and Hangzhou are all connected to Nanjing by high-speed railway and lie in the economically dynamic region of the Yangtze River Delta. The five universities are ranked in the top ten institutions for physics research in China and are especially strong in the areas of condensed-matter physics and materials sciences. CICAM brings together many important leaders in the field of artificial microstructures, including a Nobel laureate, a member of the American Academy of Engineering, 16 academicians of the Chinese Academy of Sciences, 34 Changjiang endowed professors and 45 Distinguished Young Scholars of the Natural Science Foundation of China.

Q: Why was CICAM established?

Research and development into artificial microstructures lies at the forefront of modern physical science. As one of the most important and promising research areas in the 21st century, microstructures research is at the crossover of condensed-matter physics, materials science and information science. Designing and manufacturing artificial microstructures at various scales can reveal novel quantum effects, help advance science and technology for quantum manipulation and lead to new generations of materials, information and energy technologies.



Ding Yu Xing, Co-director,
Prof. of Nanjing University.



Fu Chun Zhang, Co-director,
Prof. of Zhejiang University.



Xin Gao Gong, Co-director,
Prof. of Fudan University.

Q: What advantages do collaborative innovation centres offer?

Collaborative innovation centres attract some of the most talented researchers. They also promote interdisciplinary research through bringing together researchers with expertise in different areas and sharing resources. CICAM will combine the research capabilities of the National Laboratory of Solid State Microstructures and the State Key Laboratory of Coordination Chemistry of Nanjing University, the High Magnetic Field Laboratory at the Hefei Institutes of Physical Science of the Chinese Academy of Sciences, the State Key Laboratory of Surface Physics of Fudan University, the Key Laboratory of Artificial Structures and Quantum Control of Shanghai Jiaotong University, the Center of Correlated Matter at Zhejiang University and Huawei Technologies' Noah's Ark Lab. In addition, CICAM receives support from five provincial key laboratories and 17 national researcher training centres.

Q: What are the main research focuses of CICAM?

Focusing on cutting-edge science, CICAM chose artificial bandgap materials, correlated electron systems and small quantum systems as its three main innovation areas. It established eight cross-institutional innovation platforms: essential facilities for microstructure research, artificial bandgap- and meta-materials, micro/nano-photonics, quantum phase transitions and quantum manipulation for correlated electron

materials, novel superconducting materials and unconventional mechanisms, mesoscopic physics and devices, magnetic nanostructures and spintronics, and functional microstructured devices and system integration. Huawei participates in the construction of the last platform, which is dedicated to converting the scientific achievements of CICAM into practical applications.

“CICAM brings together many important leaders in the field of artificial microstructures”

Q: How does CICAM plan to attract talented researchers?

CICAM will follow international standards for hiring, in line with the Association of American Universities. The centre will coordinate across its different institutions, but will establish an independent management authority for establishing positions based on its research progress. The principal investigators will report directly to the centre's directors. Employment at the centre will be contract based for the duration of the research projects and key researchers will receive international evaluation. CICAM will also provide many incentives to attract talented researchers, especially active young researchers. Once hired, researchers will receive generous non-competitive research funding from the centre.



NANJING UNIVERSITY COLLABORATIVE INNOVATION CENTER OF ADVANCED MICROSTRUCTURES

WHERE MICROSTRUCTURES HAVE MACRO IMPACT

The Collaborative Innovation Center of Advanced Microstructures (CICAM), led by Nanjing University, was established under the 2011 Plan, an initiative of the Chinese Ministry of Education and Ministry of Finance to develop the innovation capacity of universities across the country. The plan was introduced as an important strategic measure in China's higher education system and through the establishment of Collaborative Innovation Centers has focused on four key categories: frontier science, industrial development, regional development and cultural heritage. In an evaluation by the Ministry of Education in early 2014, CICAM was rated top of all the Collaborative Innovation Centers considered.

CICAM concentrates on conducting interdisciplinary research on advanced and artificial microstructure materials where micro- to nano-scale features give rise to interesting properties, which can be exploited for a range of technological applications. Researchers at CICAM conduct fundamental research on advanced and artificial microstructures and also work to translate their findings into relevant applications. Through these efforts CICAM aims to establish itself as a leading centre in the field, producing research that catalyses Chinese industry and meets their technology needs. The centre

is also committed to training researchers and drawing prominent scientists from across the country to become a world-class scientific institution.

CICAM is the result of a partnership between several universities, research institutes and companies located in the Yangtze River Delta region, a thriving area for science and education in China with a very active high-tech industry. The partner institutions have a long history of cooperation in large research projects and are committed to sharing the responsibilities and benefits of CICAM.

THE STRENGTH OF COLLABORATIVE INNOVATION

CICAM's collaborative approach takes full advantage of the expertise cultivated in five established research platforms, including the National Laboratory of Solid State Microstructures at Nanjing University, to create an environment conducive to innovation on a par with that achieved by the international community.

CICAM is responsible for 60 major national research projects with a total research budget of RMB 380 million (approximately USD 62 million). Its research on dielectric superlattices and iron-based superconductors is among the strongest in the world. It has also yielded numerous important advances in the areas of optics, acoustic diodes, quantum integrated chips, high-temperature

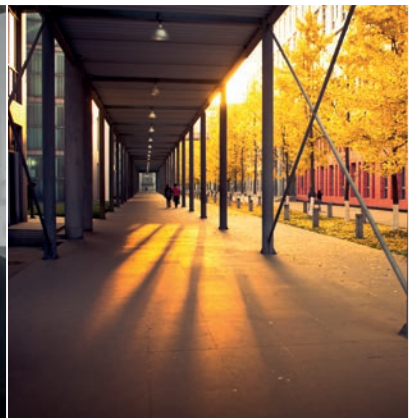
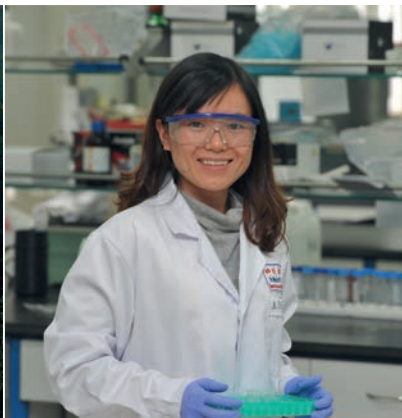
superconducting materials and their mechanisms, quantum spin Hall systems, entangled edge states, nanophotovoltaics, all-solid-state laser microstructures and micro-nanofabrication technologies. For example, CICAM researchers have designed a semiconductor laser array chip based on artificial microstructures, which entered into Huawei Technologies' industrial exploration programme on photonic integrated devices in 2013.

CICAM has attracted the attention of the scientific community for the high quality of its research as well as its implementation of novel and innovative training solutions for talented young scientists, such as personalized training for top students. An eight-year programme at the centre covering undergraduate-, masters- and PhD-level study has proved very successful.

Overall, CICAM delivers some of the best science conducted in China today. It is a dynamic centre for both research and researcher development, exploring novel approaches to artificial microstructures and producing original research to meet the country's core technology needs.



南京大學
NANJING UNIVERSITY



SICHUAN UNIVERSITY AN INNOVATION POWERHOUSE IN WESTERN CHINA

Sichuan University is located in the heart of China's Sichuan Province in the capital of Chengdu. The city has developed over the centuries into an important centre for commerce, education, transport and communication in western China.

Sichuan University is one of the oldest universities in the country, with its founding institution established as far back as 1896. The university was designated a national key university by the Chinese Ministry of Education for its excellence in education, research, and social impact, and it continues to host several key laboratories that receive financial and administrative support from the government. These include 13 national-level key laboratories and centres sponsored by the Ministry of Science and Technology, 17 national-level key laboratories and centres sponsored by the Ministry of Education and 3 national-level key laboratories sponsored by the Ministry of Health as well as numerous provincial-level key laboratories. The laboratories, research centres and bases at Sichuan University have conducted projects of regional, national and international significance.

According to Thomson Reuters' *Essential Science Indicators*, which identify the most influential researchers, publications and institutions in a range of scientific fields based on their research output and impact, Sichuan University ranks among the top one per cent globally in five subject areas, while an additional five of its disciplines are ranked among the top five per cent worldwide. Having devoted considerable resources to the areas

of teaching, learning and research, the university has gained global recognition and serves as a driver of innovation, propelling China into a new stage of economic development.

STATE KEY LABORATORY OF BIOTHERAPY

The State Key Laboratory of Biotherapy (SKLB) was founded in 2005 and selected as one of the New Drug Creation and Development Integrated Platforms in 2008 under the New Drug Creation and Development Program managed by the Ministry of Health and the Ministry of Science and Technology. In April 2013, the SKLB became the National Collaborative Innovation Center for Biotherapy, which is supported by the 2011 plan implemented by the Ministry of Education and the Ministry of Finance.

The centre's premises are divided between the medical campus of Sichuan University and the Chengdu Hi-Tech Zone. They occupy an overall area of nearly 70,000 square metres and are even now undergoing intensive growth and construction. The SKLB also takes advantage of the rich clinical resources available at the West China Hospital, Sichuan University — the largest hospital in China with 4,300 inpatient beds.

The centre excels in seamlessly integrating basic research with preclinical development and translational and clinical medicine for the discovery and development of innovative drug candidates. The establishment of an efficient and fully integrated technology chain in a single institute has proved advantageous in achieving the SKLB's ultimate goal of improving the treatment of major human diseases, including cancer, cardiovascular diseases, obesity, diabetes, inflammatory diseases, neurological diseases and chronic autoimmune diseases, as well as infectious diseases such as hepatitis, AIDS and tuberculosis.

The SKLB has almost 100 professors, associate professors and assistant professors who



四川大学
SICHUAN UNIVERSITY

LEARN MORE

Visit: www.scu.edu.cn
Address: No.24 South Section 1,
Yihuan Road,
Chengdu, China, 610065
Phone: +86-28-85468896
Fax: +86-28-85403260



are conducting well-funded, highly regarded, comprehensive and multidisciplinary research. These researchers are engaged in hundreds of projects focusing on, among other things, gene and cell therapy, vaccination, monoclonal antibodies, recombinant proteins, and the development of synthetic and natural small molecules for drug discovery. As a result of their dedicated study, the laboratory publishes over 300 research papers every year in peer-reviewed journals, including leading international journals such as the *New England Journal of Medicine*, *Developmental Cell*, *Nature Medicine*, *Proceedings of the National Academy of Sciences of the USA*, *Cancer Research* and *The Lancet Neurology*. To date, the laboratory has licensed over 50 patents in the commercial sector across China and transferred 45 potent candidate drugs to over 30 pharmaceutical companies for commercial development.

STATE KEY LABORATORY OF POLYMER MATERIALS ENGINEERING

The State Key Laboratory of Polymer Materials Engineering (SKLPME) was selected to become one of seven national pilot laboratories under the Key Discipline Development Project, which is supported by a loan from the World Bank.

The SKLPME prioritizes research at the frontier of polymer materials science and engineering that has the potential to contribute

to China's national economic development. This includes basic and applied research on the structure and properties of polymers, processing theories and related technologies, and production and engineering, in addition to the development of high-performance polymer materials.

Researchers at the SKLPME have established principles of polymer blending and compositing, developed technologies for preparing polymer-based nanomaterials and created highly efficient polymer materials for application in oil and gas fields. Researchers at the laboratory have won numerous science and technology awards, published many scientific papers and books and patented several of their innovations.

STATE KEY LABORATORY OF ORAL DISEASES

The State Key Laboratory of Oral Diseases (SKLOD) was founded in 1936 as the first research department in China specializing in oral medicine, or stomatology. It was designated a national key laboratory by the Chinese Ministry of Science and Technology.

The laboratory is primarily engaged in basic research on the mechanisms and treatment of oral diseases with the goal of becoming a leading international laboratory in the field. Research activities at the laboratory focus on developing novel techniques for the prevention and treatment of tooth decay, advancing

new dental materials and biomaterials, and understanding the mechanisms of malformation in the oral and maxillofacial area as well as the metastatic behaviour of cancerous epithelial cells that line the inside of the mouth. Researchers and postgraduate students at the SKLOD have access to the latest facilities and technologies, which cost RMB 80 million and occupy an area of 7,000 square metres.

STATE KEY LABORATORY OF HYDRAULICS AND MOUNTAIN RIVER ENGINEERING

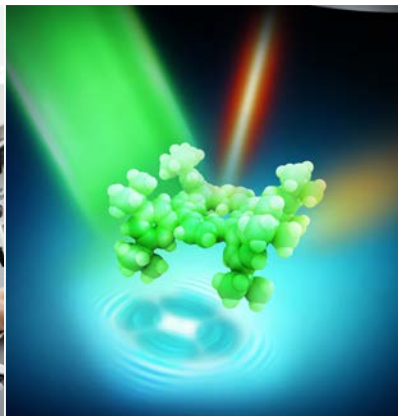
The State Key Laboratory of Hydraulics and Mountain River Engineering (SKHL) became the country's first national key laboratory in the field of hydraulic engineering, following authorization in May 1988 by the National Development and Reform Commission, formerly known as the State Planning Commission.

The laboratory was set up as an academic platform for hydraulic engineering and the study of mountain river environments to support projects in water conservation, hydropower construction and disaster prevention. The SKHL divides its research between five key objectives: the hydraulics of high-speed flow and dam engineering; mountain river dynamics and engineering; environmental hydraulics and mountain river protection; dam and reservoir safety; and hydroinformatics and new technologies in hydraulic engineering.

Between 2008 and 2012, the SKHL received one second-prize State Technological Invention Award, four second-prize State Science and Technology Progress Awards, and nine first-prize provincial- and ministerial-level allocations of the same awards. During the same period, the SKHL published 182 papers that have been included in Thomson Reuters' *Science Citation Index*, 225 papers indexed by Elsevier's *Engineering Index* and 13 monographs. The laboratory has also acquired 83 Chinese invention patents, 5 American invention patents and 8 software copyrights. Moreover, the SKHL has participated in the drafting of five volumes of technical specifications and standards.



East Gate of Jiang'an Campus



THE UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA PURSUING EXCELLENCE IN SCIENCE

The University of Science and Technology of China (USTC) is one of the most important innovation centres in the country, and is always ranked among its best universities. It is particularly strong in fields such as quantum manipulation, nanotechnology, high-temperature superconductivity, speech processing, fire science and life sciences.

The USTC takes the lead in many major science projects, such as quantum satellite research and dark-matter detection. It is also an active contributor to significant international projects, such as the International Thermonuclear Experimental Reactor (ITER) and the European Organization for Nuclear Research (CERN).

In 2013, the USTC won more than 20 renowned awards in science and technology. For example, a team of USTC physicists led by Professor Xianhui Chen received the first prize in Chinese Natural Science for their contributions to the field of superconducting materials; for the previous three years, there had been no recipients of this prize.

Some of the latest research highlights are described below.

PHYSICS AND CHEMISTRY High-energy physics at the particle colliders

A team led by Professor Zhengguo Zhao in the School of Physical Sciences made weighty contributions to the study of

diboson production, triple-gauge boson couplings and the discovery of Higgs particles via the ATLAS experiment at the Large Hadron Collider (LHC) of CERN. Zhao also greatly contributed to the observation of the Zc particles that were suggested to represent the charmed multi-quark states, using the Beijing Spectrometer (BESIII) at the Beijing Electron Positron Collider (BEPCII), and, for the first time, observed over 10 new decay modes of the charmonium states cJ and c. As a result of these outstanding achievements, Zhao was elected as an academician of the Chinese Academy of Sciences (CAS), which is the highest academic honour in the country.

Inorganic solid-state chemistry

Professor Yi Xie and her group at the Hefei National Laboratory for Physical Sciences at the Microscale (HFNL) pioneered research into the design and synthesis of inorganic functional solids with efforts to modulate their electron and phonon structures. Xie established the methodology known as the "synergetic use of binary characteristic structures" for the synthesis and assembly of inorganic functional materials, proposed a strategy for modulating the electron and phonon transport properties with phase transitions at the nanoscale, developed new high-efficiency thermoelectric materials systems, and discovered the relationship between the fine/electronic structures and the thermoelectric/optoelectronic properties of two-dimensional semiconductor crystals. As a female scientist, Xie is the youngest academician of the CAS among those elected in 2013.

Carbon aerogels sop up hydrocarbons

A team led by Professor Shuhong Yu at the HFNL is pursuing carbon aerogel



LEARN MORE

Visit: <http://en.ustc.edu.cn>
Phone: +86-(0)551-63607981



production from biomass. The team selected bacterial cellulose pellicles — a commonly used, inexpensive, nontoxic form of biomass consisting of a tangled network of cellulose nano fibres — as a precursor for the production of ultralight carbon nanofibre aerogels on a largescale. This biomass can easily be produced on an industrial scale through microbial fermentation.

QUANTUM INFORMATION AND QUANTUM TECHNOLOGY

The Synergetic Innovation Centre for Quantum Information and Quantum Physics (SIC-QIQP), head by Professor Jianwei Pan, was established and financially supported by the Chinese Ministry of Education. It focuses on bringing together teams of multidisciplinary researchers to form a dynamic national network for developing scalable quantum technologies.

Foiling quantum hackers

A research team led by Professor Qiang Zhang and Professor Tengyun Chen at the SIC-QIQP successfully demonstrated the measurement-device-independent

quantum key distribution by developing up-conversion single-photon detectors with high efficiency and low noise. The new quantum-encryption method provides the ultimate security against hackers in real-world cryptography applications, and greatly improves the security of quantum-encryption systems. This research was selected as one of the Highlights of the Year in *Physics* by the American Physical Society.

A milestone in satellite-based quantum communication

A collaborative team led by Professor Chengzhi Peng at the SIC-QIQP achieved comprehensive and direct verification of quantum communication between satellites and ground stations. This research lays the necessary technical foundations for a global quantum-communication network based on ground-satellite quantum communication by launching the quantum science experimental satellite of China.

Optical spectroscopy goes intramolecular

A team led by Professor Zhenchao Dong at the SIC-QIQP reported an optical spectroscopic-imaging approach that achieves subnanometre resolution and resolves the internal structure of single molecules. This development could lead to new techniques for probing and controlling nanoscale structure, dynamics, mechanics and chemistry. This research was listed among China's top 10 science news stories in 2013.

ENVIRONMENTAL AND EARTH SCIENCES

Penguins thrived in Antarctica during the Little Ice Age

New research led by Professor Liguang Sun in the School of Earth and Space Sciences showed that penguin populations in the Ross Sea of Antarctica spiked during the short cold period, called the Little Ice Age, which occurred between AD1500 and 1800. These results run contrary to previous studies that found increases in Antarctic penguin populations during warmer periods and decreases during

colder periods, suggesting that populations living at different latitudes in the Antarctic might respond differently to climate change.

Uncovering the mystery of subduction zone earthquakes

Based on analytical data from four of the highest magnitude subduction zone megathrust earthquakes, the conclusion was drawn that low-frequency radiation is closer to the trench at shallower depths and high-frequency radiation is farther from the trench at greater depths, in general. This scientific breakthrough was achieved by a team led by Professor Huajian Yao.

LIFE SCIENCES

New evidence for curing type 2 diabetes

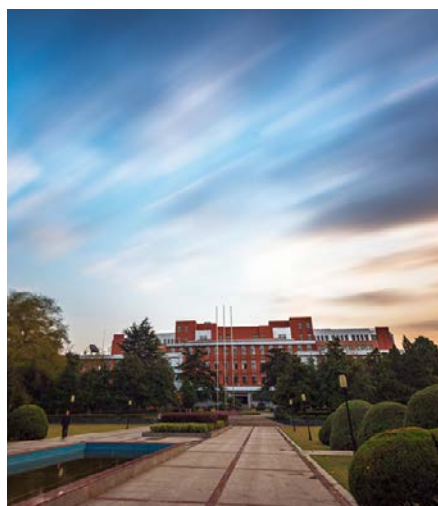
Research teams led by Professor Rongbin Zhou and Professor Zhigang Tian in the School of Life Sciences revealed a new mechanism through which omega-3 fatty acids inhibit inflammation and prevent type 2 diabetes. The research results were published in *Immunity* in June 2013 and highlighted in the same issue of the journal.

Identifying liver-resident natural-killer cells with immune memory

A team also led by Professor Zhigang Tian identified liver-resident natural-killer (NK) cells that possess unique immune memory characteristics absent from normal NK cells.

LincRNA-p21 as a novel key player in regulating the Warburg effect

A research team led by Professor Mian Wu and Professor Yide Mei, at HFNL and the School of Life Sciences, has revealed a novel mechanism whereby lincRNA-p21 regulates the Warburg effect under hypoxic conditions. They demonstrated, for the first time, that lincRNA-p21 is an important regulator of the Warburg effect, and also identify lincRNA-p21 as a valuable therapeutic target for cancer.



Guo Moruo Square. USTC was established by the Chinese Academy of Sciences (CAS) in 1958 in Beijing. The director of CAS, Mr. Guo Moruo was appointed the first president of USTC.